

Smart Antennas for Wireless Communications

With MATLAB

FRANK GROSS

**Smart Antennas
for Wireless
Communications**

This page intentionally left blank

Smart Antennas for Wireless Communications

With MATLAB

Frank B. Gross, PhD

Senior Systems Engineer

Argon ST

Fairfax, Virginia

McGraw-Hill

New York Chicago San Francisco Lisbon London Madrid
Mexico City Milan New Delhi San Juan Seoul
Singapore Sydney Toronto

Copyright © 2005 by The McGraw-Hill Companies, Inc. All rights reserved. Manufactured in the United States of America. Except as permitted under the United States Copyright Act of 1976, no part of this publication may be reproduced or distributed in any form or by any means, or stored in a database or retrieval system, without the prior written permission of the publisher.

0-07-158895-7

The material in this eBook also appears in the print version of this title: 0-07-144789-X.

All trademarks are trademarks of their respective owners. Rather than put a trademark symbol after every occurrence of a trademarked name, we use names in an editorial fashion only, and to the benefit of the trademark owner, with no intention of infringement of the trademark. Where such designations appear in this book, they have been printed with initial caps.

McGraw-Hill eBooks are available at special quantity discounts to use as premiums and sales promotions, or for use in corporate training programs. For more information, please contact George Hoare, Special Sales, at george_hoare@mcgraw-hill.com or (212) 904-4069.

TERMS OF USE

This is a copyrighted work and The McGraw-Hill Companies, Inc. ("McGraw-Hill") and its licensors reserve all rights in and to the work. Use of this work is subject to these terms. Except as permitted under the Copyright Act of 1976 and the right to store and retrieve one copy of the work, you may not decompile, disassemble, reverse engineer, reproduce, modify, create derivative works based upon, transmit, distribute, disseminate, sell, publish or sublicense the work or any part of it without McGraw-Hill's prior consent. You may use the work for your own noncommercial and personal use; any other use of the work is strictly prohibited. Your right to use the work may be terminated if you fail to comply with these terms.

THE WORK IS PROVIDED "AS IS." McGRAW-HILL AND ITS LICENSORS MAKE NO GUARANTEES OR WARRANTIES AS TO THE ACCURACY, ADEQUACY OR COMPLETENESS OF OR RESULTS TO BE OBTAINED FROM USING THE WORK, INCLUDING ANY INFORMATION THAT CAN BE ACCESSED THROUGH THE WORK VIA HYPERLINK OR OTHERWISE, AND EXPRESSLY DISCLAIM ANY WARRANTY, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. McGraw-Hill and its licensors do not warrant or guarantee that the functions contained in the work will meet your requirements or that its operation will be uninterrupted or error free. Neither McGraw-Hill nor its licensors shall be liable to you or anyone else for any inaccuracy, error or omission, regardless of cause, in the work or for any damages resulting therefrom. McGraw-Hill has no responsibility for the content of any information accessed through the work. Under no circumstances shall McGraw-Hill and/or its licensors be liable for any indirect, incidental, special, punitive, consequential or similar damages that result from the use of or inability to use the work, even if any of them has been advised of the possibility of such damages. This limitation of liability shall apply to any claim or cause whatsoever whether such claim or cause arises in contract, tort or otherwise.

DOI: 10.1036/007144789X



Professional



Want to learn more?

We hope you enjoy this McGraw-Hill eBook! If you'd like more information about this book, its author, or related books and websites, please [click here](#).

Contents at a Glance

Preface	xi	
1	Introduction	1
2	Fundamentals of Electromagnetic Fields	9
3	Antenna Fundamentals	37
4	Array Fundamentals	65
5	Principles of Random Variables and Processes	105
6	Propagation Channel Characteristics	123
7	Angle-of-Arrival Estimation	169
8	Smart Antennas	207
Index	267	

This page intentionally left blank

Contents

Preface	xi
---------	----

Chapter 1. Introduction	1
1.1 What is a Smart Antenna?	1
1.2 Why are Smart Antennas Emerging Now?	2
1.3 What are the Benefits of Smart Antennas?	3
1.4 Smart Antennas Involve Many Disciplines	5
1.5 Overview of the Book	6
References	7

Chapter 2. Fundamentals of Electromagnetic Fields	9
2.1 Maxwell's Equations	9
2.2 The Helmholtz Wave Equation	11
2.3 Propagation in Rectangular Coordinates	12
2.4 Propagation in Spherical Coordinates	14
2.5 Electric Field Boundary Conditions	15
2.6 Magnetic Field Boundary Conditions	19
2.7 Planewave Reflection and Transmission Coefficients	21
2.7.1 Normal incidence	21
2.7.2 Oblique incidence	24
2.8 Propagation Over Flat Earth	27
2.9 Knife-Edge Diffraction	31
References	33
Problems	33

Chapter 3. Antenna Fundamentals	37
3.1 Antenna Field Regions	37
3.2 Power Density	39
3.3 Radiation Intensity	42
3.4 Basic Antenna Nomenclature	44
3.4.1 Antenna pattern	44
3.4.2 Antenna boresight	46

viii Contents

3.4.3	Principal plane patterns	46
3.4.4	Beamwidth	47
3.4.5	Directivity	48
3.4.6	Beam solid angle	49
3.4.7	Gain	49
3.4.8	Effective aperture	49
3.5	Friis Transmission Formula	50
3.6	Magnetic Vector Potential and the Far Field	51
3.7	Linear Antennas	53
3.7.1	Infinitesimal dipole	53
3.7.2	Finite length dipole	55
3.8	Loop Antennas	58
3.8.1	Loop of constant phasor current	58
	References	61
	Problems	61
Chapter 4.	Array Fundamentals	65
4.1	Linear Arrays	65
4.1.1	Two element array	66
4.1.2	Uniform N -element linear array	68
4.1.3	Uniform N -element linear array directivity	76
4.2	Array Weighting	79
4.2.1	Beamsteered and weighted arrays	88
4.3	Circular Arrays	89
4.3.1	Beamsteered circular arrays	90
4.4	Rectangular Planar Arrays	91
4.5	Fixed Beam Arrays	93
4.5.1	Butler matrices	94
4.6	Fixed Sidelobe Canceling	95
4.7	Retrodirective Arrays	98
4.7.1	Passive retrodirective array	99
4.7.2	Active retrodirective array	100
	References	101
	Problems	102
Chapter 5.	Principles of Random Variables and Processes	105
5.1	Definition of Random Variables	105
5.2	Probability Density Functions	106
5.3	Expectation and Moments	108
5.4	Common Probability Density Functions	109
5.4.1	Gaussian density	110
5.4.2	Rayleigh density	111
5.4.3	Uniform density	111
5.4.4	Exponential density	113
5.4.5	Rician density	114
5.4.6	Laplace density	115
5.5	Stationarity and Ergodicity	115
5.6	Autocorrelation and Power Spectral Density	117
5.7	Correlation Matrix	119

References	120
Problems	120
Chapter 6. Propagation Channel Characteristics	123
6.1 Flat Earth Model	124
6.2 Multipath Propagation Mechanisms	127
6.3 Propagation Channel Basics	129
6.3.1 Fading	129
6.3.2 Fast fading modeling	130
6.3.3 Channel impulse response	141
6.3.4 Power delay profile	142
6.3.5 Prediction of power delay profiles	144
6.3.6 Power angular profile	145
6.3.7 Prediction of angular spread	148
6.3.8 Power delay-angular profile	151
6.3.9 Channel dispersion	151
6.3.10 Slow fading modeling	153
6.4 Improving Signal Quality	155
6.4.1 Equalization	156
6.4.2 Diversity	158
6.4.3 Channel coding	159
6.4.4 MIMO	160
References	163
Problems	165
Chapter 7. Angle-of-Arrival Estimation	169
7.1 Fundamentals of Matrix Algebra	169
7.1.1 Vector basics	170
7.1.2 Matrix basics	171
7.2 Array Correlation Matrix	175
7.3 AOA Estimation Methods	178
7.3.1 Bartlett AOA estimate	178
7.3.2 Capon AOA estimate	179
7.3.3 Linear prediction AOA estimate	181
7.3.4 Maximum entropy AOA estimate	182
7.3.5 Pisarenko harmonic decomposition AOA estimate	183
7.3.6 Min-norm AOA estimate	185
7.3.7 MUSIC AOA estimate	187
7.3.8 Root-MUSIC AOA estimate	191
7.3.9 ESPRIT AOA estimate	197
References	201
Problems	202
Chapter 8. Smart Antennas	207
8.1 Introduction	207
8.2 The Historical Development of <i>Smart Antennas</i>	209
8.3 Fixed Weight Beamforming Basics	211
8.3.1 Maximum signal-to-interference ratio	211
8.3.2 Minimum mean-square error	218

x Contents

8.3.3	Maximum likelihood	221
8.3.4	Minimum variance	223
8.4	Adaptive Beamforming	227
8.4.1	Least mean squares	227
8.4.2	Sample matrix inversion	230
8.4.3	Recursive least squares	234
8.4.4	Constant modulus	238
8.4.5	Least squares constant modulus	240
8.4.6	Conjugate gradient method	246
8.4.7	Spreading sequence array weights	250
8.4.8	Description of the new SDMA receiver	252
	References	260
	Problems	262

Index	267
-------	-----

Preface

The subject of smart antennas is beginning to enjoy immense popularity due to the current exponential growth in all forms of wireless communications and sensing. This rapid growth has been facilitated by advancing digital signal processing hardware and also by the global interest in wideband wireless applications. Many in the wireless communication world are hoping to utilize smart antennas to boost capacities, expand bandwidths, increase signal-to-interference ratios, mitigate fading, and improve MIMO communications. Many others in the defense community are hoping to utilize smart antennas to help facilitate secure communications, direction finding, waveform diversity applications, MIMO radar, and multi-mission operations. Both groups are looking to use the benefits of smart antennas for similar reasons but are approaching the subject from different requirements and perspectives. This text has been written in the hopes of providing each group with a fundamental understanding of smart antennas without needing to tie the treatment to any one specific application.

This author has an extensive background in signal processing, radar, communications, and electromagnetics. Having worked in industry and in academia, it has been difficult to find a smart antenna text that can be equally shared by both communities. In addition, only a few books in print even address the subject of smart antennas. Since smart antennas involve an amalgamation of many different disciplines, a background in each related area must be understood in order to appreciate this topic as a whole. Thus, one overriding goal of this text is to present the fundamentals of several different science and engineering disciplines. The intent is to show how all of these disciplines converge in the study of smart antennas. To understand smart antenna behavior, one must be versed in various topics such as electromagnetics, antennas, array processing, propagation, channel characterization, random processes, spectral estimation, and adaptive methods. Thus, the book lays a background in each of these disciplines before tackling smart antennas in particular.

This text is organized into eight chapters. Chapter 1 gives the background, motivation, justification, and benefits for the study of this important topic. Chapter 2 provides a summary of electromagnetics, reflection, diffraction, and propagation. These concepts help in understanding the path gain factor, coverage diagrams, and fading. This material will be used later in order to understand the nature of multipath propagation and the phase relationship between different array elements. Chapter 3 deals with general antenna theory including antenna metrics such as beamwidth, gain, principle plane patterns, and effective apertures. The Friis transmission formula is discussed because it aids in the understanding of spherical spreading and reception issues. Finally the specific behavior of dipoles and loops is addressed. The goal is to help the reader gain a basic understanding of how individual antenna elements affect the behavior of arrays. Chapter 4 addresses the subject of antenna arrays. Array phenomenology is addressed in order to help in understanding the relationship of array shape to beam patterns. Linear, circular, and planar arrays are discussed. Array weighting or “shading” is explored in order to help the reader understand how array weights influence the radiation pattern. Specific arrays are discussed such as fixed beam arrays, beamsteered arrays, Butler matrices, and retrodirective arrays. This treatment of array behavior is invaluable in understanding smart antenna behavior and limitations. Chapter 5 lays a foundation in random variables and processes. This is necessary because multipath signals and noise are characterized by random behavior. Also, channel delays and angles of arrival tend to be random variables. Thus, a minimum background in random processes must be established in order to understand the nature of arriving signals and how to process array inputs. Many smart antenna applications require the computation of the array correlation matrix. The topics of ergodicity and stationarity are discussed to help the reader understand the nature of the correlation matrix, prediction of the matrix, and what information can be extracted in order to compute the optimum array weights. It is assumed that students taking a smart antenna class are already familiar with random processes. However, this chapter is provided to help understand gird concepts, which will be addressed in later chapters. Chapter 6 addresses propagation channel characteristics. Such critical issues as fading, delay spread, angular spread, dispersion, and equalization are discussed. In addition, MIMO is briefly defined and addressed. If one understands the nature of multipath fading, one can better design a smart antenna that minimizes the deleterious effects. Chapter 7 discusses numerous different spectral estimation methods. The topics range from Bartlett beamforming to Pisarenko harmonic decomposition to eigenstructure methods such as MUSIC and ESPRIT. This chapter helps in understanding many useful properties of the

array correlation matrix as well as demonstrating that angles of arrival can be predicted with greater accuracy than the array resolution can allow. In addition, many of the techniques discussed in Chapter 7 help aid in understanding adaptive methods. Chapter 8 shows the historical development of smart antennas and how weights can be computed by minimizing cost functions. The minimum mean-squared-error method lends itself to understanding iterative methods such as least mean squares. Several iterative methods are developed and discussed and their performance is contrasted with numerous examples. Current popular methods such as the constant modulus algorithm, sample matrix inversion, and the conjugate gradient method are explored. Lastly, a waveform diversity concept is discussed wherein a different waveform is applied to each array element in order to determine angles of arrival. This method has application to both MIMO communications and MIMO radar.

Numerous MATLAB examples are given in the text, and most homework questions require the use of MATLAB. It is felt that if students can program smart antenna algorithms in MATLAB, a further depth of understanding can be achieved. All MATLAB codes written and used for the completion of the book are available for use by the reader. It is intended that these codes can serve as templates for further work.

Examples and Solutions on the Web

The MATLAB codes are available at the website: *<http://books.mcgraw-hill.com>*. The codes are organized into three categories—examples, figures, and problems. There are codes produced in conjunction with example problems in most chapters. The codes are denoted by sa_ex#.#.m. For example, Example 8.4 in Chapter 8 will have an associated MATLAB code labeled sa_ex8_4.m. There are codes used to produce most of the figures in the book. These codes are denoted by sa_fig#.#.m. Thus, if one is interested in replicating a book figure or modifying a book figure, the code associated with that figure can be downloaded and used. There are also codes used to produce most of the homework solutions. These codes are denoted by sa_prob#.#.m. The students will have download access to the codes associated with the figures and examples. The instructor will have access to the codes associated with figures, examples, and problem solutions.

For Instructors

This book can be used as a one-semester graduate or advanced undergraduate text. The instructor should allot two to three weeks for exploring Chapter 7 and three to four weeks for exploring Chapter 8.

The extent to which earlier chapters are treated depends upon the background of the students. It is normally assumed that the prerequisites for the course are an undergraduate course in communications, an undergraduate course in advanced electromagnetics, and a basic course in random processes. The students may be able to take this course if they have a background in random processes, which is addressed in an undergraduate course in communications. It is assumed that all students have had an advanced engineering math course, which covers matrix theory including the calculation of eigenvalues and eigenvectors.

It is my hope that this book will open new doors of understanding for the uninitiated and serve as a good resource material for the practitioner.

Acknowledgments

This author wishes to acknowledge the tremendous help and support of many individuals. A special thanks goes to the very supportive staff at McGraw-Hill. The editor Steve Chapman has always believed in this project and has provided me with continuous help in completing the work. Great appreciation is also given to Wendy Rinaldi and Diana Mattingly who have helped encourage the progress, find reviewers, and collate the manuscript. I am thankful for the expert review and helpful comments from Assistant Professor Fernando Teixeira of the Ohio State University. His comments and insights have helped strengthen the book. I am appreciative of my EEL 5930 Smart Antennas class. The book was road tested on them and they happily survived the experience. A deep debt of gratitude is given to Jeff Connor, my graduate student. He performed outstanding work in helping to proof the manuscript, solve all of the homework problems, and organize and edit all of the MATLAB script files. He worked tirelessly for long hours in providing both excellent and aesthetic supporting materials. His enthusiasm for this project was an added inspiration.

**Smart Antennas
for Wireless
Communications**

This page intentionally left blank

Introduction

This text has been written in response to the recent extreme interest in the rapidly growing field of smart antennas. Although some of the principles of smart antennas have been around for over forty years, new wireless applications demanding smart antenna technology are growing exponentially. In addition, the latest algorithms that control smart antennas have matured to the point of being extremely effective in dynamic and dispersive multipath environments. Thus, smart antennas are now becoming a critical adjunct for increasing the performance of a myriad of wireless applications. This new technology has a major role in all forms of wireless systems ranging from mobile cellular to *personal communications services* (PCS) to radar. This text will not address specific applications as much as it will introduce the reader to the basic principles which underlie smart antennas. A solid foundation is necessary in order to understand the full applicability and benefit of this rapidly growing technology.

1.1 What is a Smart Antenna?

The term “smart antenna” generally refers to any antenna array, terminated in a sophisticated signal processor, which can adjust or adapt its own beam pattern in order to emphasize signals of interest and to minimize interfering signals.

Smart antennas generally encompass both switched beam and beamformed adaptive systems. Switched beam systems have several available fixed beam patterns. A decision is made as to which beam to access, at any given point in time, based upon the requirements of the system. Beamformed adaptive systems allow the antenna to steer the beam to any direction of interest while simultaneously nulling interfering signals. The smart antenna concept is opposed to the fixed

2 Chapter One

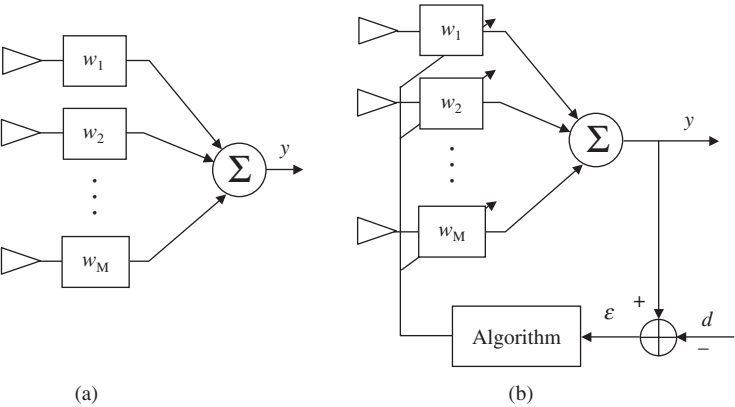


Figure 1.1 (a) Traditional array (b) Smart antenna.

beam “dumb antenna,” which does not attempt to adapt its radiation pattern to an ever-changing electromagnetic environment. In the past, smart antennas have alternatively been labeled adaptive arrays or digital beamforming arrays. This new terminology reflects our penchant for “smart” technologies and more accurately identifies an adaptive array that is controlled by sophisticated signal processing. Figure 1.1 contrasts two antenna arrays. The first is a traditional, fixed beam array where the mainlobe can be steered, by defining the fixed array weights $\bar{w}$. However, this configuration is neither smart nor adaptive.

The second array in the figure is a smart antenna designed to adapt to a changing signal environment in order to optimize a given algorithm. An optimizing criterion, or cost function, is normally defined based upon the requirements at hand. In this example, the cost function is defined as the *magnitude of the error squared*, $|\varepsilon|^2$, between the desired signal d and the array output y . The array weights $\bar{w}$ are adjusted until the output matches the desired signal and the cost function is minimized. This results in an optimum radiation pattern.

1.2 Why are Smart Antennas Emerging Now?

The rapid growth in demand for smart antennas is fueled by two major reasons. First, the technology for high speed *analog-to-digital converters* (ADC) and high speed digital signal processing is burgeoning at an alarming rate. Even though the concept of smart antennas has been around since the late 50s [1–3], the technology required in order to make the necessary rapid and computationally intense calculations has only emerged recently. Early smart antennas, or adaptive arrays, were

limited in their capabilities because adaptive algorithms were usually implemented in analog hardware. With the growth of ADC and *digital signal processing* (DSP); what was once performed in hardware can now be performed digitally and quickly [4]. ADCs, which have resolutions that range from 8 to 24 bits, and sampling rates approaching 20 *Gigasamples per second* (GSa/s), are now a reality [5]. In time, superconducting data converters will be able to sample data at rates up to 100 GSa/s [6]. This makes the direct digitization of most *radio frequency* (RF) signals possible in many wireless applications. At the very least, ADC can be applied to IF frequencies in higher RF frequency applications. This allows most of the signal processing to be defined in software near the front end of the receiver. In addition, DSP can be implemented with high speed parallel processing using *field programmable gate arrays* (FPGA). Current commercially available FPGAs have speeds of up to 256 BMACS.¹ Thus, the benefits of smart antenna integration will only flourish, given the exponential growth in the enabling digital technology continues.

Second, the global demand for all forms of wireless communication and sensing continues to grow at a rapid rate. Smart antennas are the practical realization of the subject of adaptive array signal processing and have a wide range of interesting applications. These applications include, but are not limited to, the following: mobile wireless communications [7], software-defined radio [8, 9], *wireless local area networks* (WLAN) [10], *wireless local loops* (WLL) [11], mobile Internet, *wireless metropolitan area networks* (WMAN) [12], satellite-based personal communications services, radar [13], ubiquitous radar [14], many forms of remote sensing, *mobile ad hoc networks* (MANET) [15], high data rate communications [16], satellite communications [17], *multiple-in-multiple-out* (MIMO) systems [18], and waveform diversity systems [19].

The rapid growth in telecommunications alone is sufficient to justify the incorporation of smart antennas to enable higher system capacities and data rates. It is projected that the United States will spend over \$137 billion on telecommunications in the year 2006. Global expenditures on telecommunications are rapidly approaching \$3 trillion.

1.3 What are the Benefits of Smart Antennas?

Smart antennas have numerous important benefits in wireless applications as well as in sensors such as radar. In the realm of mobile wireless applications, smart antennas can provide higher system capacities by

¹BMACS: Billion multiply accumulates per second.

4 Chapter One

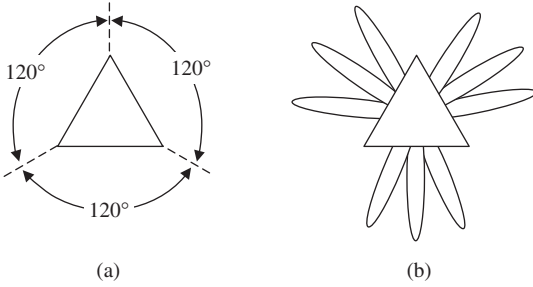


Figure 1.2 (a) Sectorized (b) Smart Antennas.

directing narrow beams toward the users of interest, while nulling other users not of interest. This allows for higher signal-to-interference ratios, lower power levels, and permits greater frequency reuse within the same cell. This concept is called *space division multiple access* (SDMA). In the United States, most base stations sectorize each cell into three 120° swaths as seen in Fig. 1.2a. This allows the system capacity to potentially triple within a single cell because users in each of the three sectors can share the same spectral resources. Most base stations can be modified to include smart antennas within each sector. Thus the 120° sectors can be further subdivided as shown in Fig. 1.2b. This further subdivision enables the use of lower power levels, and provides for even higher system capacities and greater bandwidths.

Another benefit of smart antennas is that the deleterious effects of multipath can be mitigated. As will be discussed in Chap. 8, a constant modulus algorithm, which controls the smart antenna, can be implemented in order to null multipath signals. This will dramatically reduce fading in the received signal. Higher data rates can be realized because smart antennas can simultaneously reduce both co-channel interference and multipath fading. Multipath reduction not only benefits mobile communications but also applies to many applications of radar systems.

Smart antennas can be used to enhance *direction-finding* (DF) techniques by more accurately finding *angles-of-arrival* (AOA) [20]. A vast array of spectral estimation techniques can be incorporated, which are able to isolate the AOA with an angular precision that exceeds the resolution of the array. This topic will be discussed in detail in Chap. 7. The accurate estimation of AOA is especially beneficial in radar systems for imaging objects or accurately tracking moving objects. Smart antenna DF capabilities also enhance geo-location services enabling a wireless system to better determine the location of a particular mobile user. Additionally, smart antennas can direct the array main beam toward

signals of interest even when no reference signal or training sequence is available. This capability is called blind adaptive beamforming.

Smart antennas also play a role in MIMO communications systems [18] and in waveform diverse MIMO radar systems [21, 22]. Since diverse waveforms are transmitted from each element in the transmit array and are combined at the receive array, smart antennas will play a role in modifying radiation patterns in order to best capitalize on the presence of multipath. With MIMO radar, the smart antenna can exploit the independence between the various signals at each array element in order to use target scintillation for improved performance, to increase array resolution, and to mitigate clutter [19].

Many smart antenna benefits will be discussed in detail in Chaps. 7 and 8. In summary, let us list some of the numerous potential benefits of smart antennas.

- Improved system capacities
- Higher permissible signal bandwidths
- Space division multiple access (SDMA)
- Higher signal-to-interference ratios
- Increased frequency reuse
- Sidelobe canceling or null steering
- Multipath mitigation
- Constant modulus restoration to phase modulated signals
- Blind adaptation
- Improved angle-of-arrival estimation and direction finding
- Instantaneous tracking of moving sources
- Reduced speckle in radar imaging
- Clutter suppression
- Increased degrees of freedom
- Improved array resolution
- MIMO compatibility in both communications and radar

1.4 Smart Antennas Involve Many Disciplines

The general subject of smart antennas is the necessary union between such related topics as electromagnetics, antennas, propagation, communications, random processes, adaptive theory, spectral estimation, and array signal processing. Figure 1.3 demonstrates the important

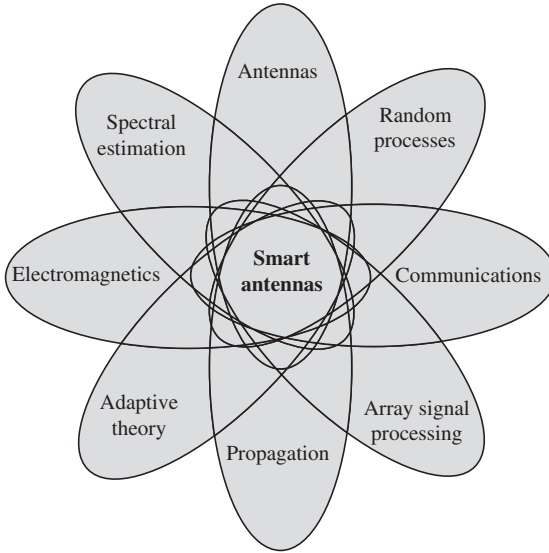


Figure 1.3 Venn diagram relating various disciplines to smart antennas.

relationship between each discipline. Many previous attempts have been made to explain smart antennas from the background of a single discipline; however, this myopic approach appeals only to small segments of the engineering community and does not yield a full appreciation for this valuable subject. No single engineering discipline can be the sole province of this rapidly growing field. The subject of smart antennas transcends specific applications and thus merits a more global treatment. In order to fundamentally understand smart antennas, one must be versed in many varied and related topics. One could argue that some of the disciplines displayed in Fig. 1.3 can be merged to create a smaller list. However, the specialist, in each of these specific disciplines, brings a unique contribution to the general field of smart antennas. Thus, this book is an attempt to preview all of the above disciplines and to relate each of them to the subject as a whole.

1.5 Overview of the Book

As has been mentioned, this text has been written in order to provide the reader with a broad and fundamental understanding of the subject of smart antennas. A foundation of basic principles is laid in each of the supporting disciplines which relate to smart antennas as a whole. These various disciplines are outlined chapter by chapter in this book. Electromagnetics, antennas, and arrays are discussed in Chaps. 2 to 4. This

foundation is critical in better understanding the physics supporting smart antenna behavior. Random processes along with specific probability distributions are discussed in Chap. 5. This aids in the understanding of noise, channels, delay spread, angular spread, and channel profiles. Propagation channel characterization is discussed in Chap. 6. This helps the reader to understand the basics of multipath and fading. This will prove invaluable in understanding the limitations of smart antenna performance. AOA estimation is discussed in Chap. 7. This chapter explores the various techniques employed to estimate AOA and lays a foundation for understanding eigenstructure methods. Finally, smart antennas are discussed at length in Chap. 8. The history of smart antenna development is discussed. Numerous adaptive algorithms are explained and explored.

The intent is that the interested reader can use the principles outlined in this book in order to become very knowledgeable in the fundamentals of smart antennas. In addition, numerous MATLAB examples are given. It is believed that a concept can be more fully understood if it can be modeled in software and solutions can be visualized. Numerous MATLAB script files are provided on a CD so that the student can understand how to program these various algorithms and observe their performance. It is ultimately hoped, with this fundamental understanding, that the student can use this information as a spring-board for more advanced work in this fascinating field. The chapters are summarized as follows:

Chapter 1: Introduction

Chapter 2: Fundamentals of Electromagnetic Fields

Chapter 3: Antenna Fundamentals

Chapter 4: Array Fundamentals

Chapter 5: Principles of Random Variables and Processes

Chapter 6: Propagation Channel Characteristics

Chapter 7: Angle-of-Arrival Estimation

Chapter 8: Smart Antennas

References

1. Van Atta, L. "Electromagnetic Reflection," U.S. Patent 2908002, Oct. 6, 1959.
2. Howells, P. "Intermediate Frequency Sidelobe Canceller," U.S. Patent 3202990, Aug. 24, 1965.
3. Howells, P. "Explorations in Fixed and Adaptive Resolution at GE and SURC," *IEEE Transactions on Antenna and Propagation*, Special Issue on Adaptive Antennas, Vol. AP-24, No. 5, pp. 575–584, Sept. 1976.
4. Walden, R.H. "Performance Trends for Analog-to-Digital Converters," *IEEE Commn. Mag.*, pp. 96–101, Feb. 1999.

8 Chapter One

5. Litva, J., and T. Kwok-Yeung Lo, *Digital Beamforming in Wireless Communications*, Artech House, Boston, MA, 1996.
6. Brock, D.K., O.A. Mukhanov, and J. Rosa, "Superconductor Digital RF Development for Software Radio," *IEEE Commun. Mag.*, pp. 174, Feb. 2001.
7. Liberti, J., and T. Rappaport, *Smart Antennas for Wireless Communications: IS-95 and Third Generation CDMA Applications*, Prentice Hall New York, 1999.
8. Reed, J., *Software Radio: A Modern Approach to Radio Engineering*, Prentice Hall, New York, 2002.
9. Mitola, J., "Software Radios," *IEEE Commun. Mag.*, May 1995.
10. Doufexi, A., S. Armour, A. Nix, P. Karlsson, and D. Bull, "Range and Throughput Enhancement of Wireless Local Area Networks Using Smart Sectorised Antennas," *IEEE Transactions on Wireless Communications*, Vol. 3, No. 5, pp. 1437–1443, Sept. 2004.
11. Weisman, C., *The Essential Guide to RF and Wireless*, 2d ed., Prentice Hall, New York 2002.
12. Stallings, W., *Local and Metropolitan Area Networks*, 6th ed., Prentice Hall, New York, 2000.
13. Skolnik, M., *Introduction to Radar Systems*, McGraw-Hill, 3d ed., New York, 2001.
14. Skolnik, M., "Attributes of the Ubiquitous Phased Array Radar," *IEEE Phased Array Systems and Technology Symposium*, Boston, MA, Oct. 14–17, 2003.
15. Lal, D., T. Joshi, and D. Agrawal, "Localized Transmission Scheduling for Spatial Multiplexing Using Smart Antennas in Wireless Adhoc Networks," *13th IEEE Workshop on Local and Metropolitan Area Networks*, pp. 175–180, April 2004.
16. Wang Y., and H. Scheving, "Adaptive Arrays for High Rate Data Communications," *48th IEEE Vehicular Technology Conference*, Vol. 2, pp. 1029–1033, May 1998.
17. Jeng, S., and H. Lin, "Smart Antenna System and Its Application in Low-Earth-Orbit Satellite Communication Systems," *IEE Proceedings on Microwaves, Antennas, and Propagation*, Vol. 146, No. 2, pp. 125–130, April 1999.
18. Durgin, G., *Space-Time Wireless Channels*, Prentice Hall, New York, 2003.
19. Ertan, S., H. Griffiths, M. Wicks, et al., "Bistatic Radar Denial by Spatial Waveform Diversity," *IEE RADAR 2002*, Edinburgh, pp. 17–21, Oct. 15–17, 2002.
20. Talwar, S., M. Viberg, and A. Paulraj, "Blind Estimation of Multiple Co-Channel Digital Signals Using an Atnenna Array," *IEEE Signal Processing Letters*, Vol. 1, No. 2, Feb. 1994.
21. Rabideau, D., and P. Parker, "Ubiquitous MIMO Multifunction Digital Array Radar," *IEEE Signals, Systems, and Computers*, 37th Asilomar Conference, Vol. 1, pp. 1057–1064, Nov. 9–12, 2003.
22. Fishler, E., A. Haimovich, R. Blum, et al., "MIMO Radar: An Idea Whose Time Has Come," *Proceedings of the IEEE Radar Conference*, pp. 71–78, April 26–29, 2004.

Fundamentals of Electromagnetic Fields

The foundation for all wireless communications is based upon understanding the radiation and reception of wireless antennas as well as the propagation of electromagnetic fields between these antennas. Regardless of the form of wireless communications used or the particular modulation scheme chosen, wireless communication is based upon the laws of physics. Radiation, propagation, and reception can be explained through the use of Maxwell's four foundational equations.

2.1 Maxwell's Equations

It was the genius of James Clerk Maxwell¹ to combine the previous work of Michael Faraday,² Andre Marie Ampere,³ and Carl Fredrick Gauss⁴ into one unified electromagnetic theory. (Some very useful references describing electromagnetics basics are Sadiku [1], Hayt [2], and

¹James Clerk Maxwell (1831–1879): A Scottish born physicist who published his treatise on electricity and magnetism in 1873.

²Michael Faraday (1791–1867): An English born chemist and experimenter who connected time varying magnetic fields with induced currents.

³Andre Marie Ampere (1775–1836): A French born physicist who found that current in one wire exerts force on another wire.

⁴Carl Fredrick Gauss (1777–1855): A German born mathematical genius who helped to establish a worldwide network of terrestrial magnetism observation points.

Ulaby [3].) Maxwell's equations are given as follows:

$$\text{Faraday's law} \quad \nabla \times \bar{E} = -\frac{\partial \bar{B}}{\partial t} \quad (2.1)$$

$$\text{Ampere's law} \quad \nabla \times \bar{H} = \frac{\partial \bar{D}}{\partial t} + \bar{J} \quad (2.2)$$

$$\text{Gauss's laws} \quad \begin{cases} \nabla \cdot \bar{D} = \rho \\ \nabla \cdot \bar{B} = 0 \end{cases} \quad (2.3)$$

$$(2.4)$$

where $\bar{E}$ = electric field intensity vector (V/m)

$\bar{D}$ = electric flux density vector (C/m²)

$\bar{H}$ = magnetic field intensity vector (A/m)

$\bar{B}$ = magnetic flux density vector (W/m²)

$\bar{J}$ = volume current density vector (A/m²)

ρ = volume charge density (C/m³)

The electric flux density and the electric field intensity are related through the permittivity of the medium as given by

$$\bar{D} = \varepsilon \bar{E} \quad (2.5)$$

The magnetic flux density and the magnetic field intensity are related through the permeability of the medium as given by

$$\bar{B} = \mu \bar{H} \quad (2.6)$$

where $\varepsilon = \varepsilon_r \varepsilon_0$ = permittivity of the medium (F/m)

ε_0 = permittivity of free space = 8.85×10^{-12} F/m

$\mu = \mu_r \mu_0$ = permeability of the medium (H/m)

μ_0 = permeability of free space = $4\pi \times 10^{-7}$ H/m

With no sources present and expressing the fields as the phasors $\bar{E}_s$ and $\bar{H}_s$, Maxwell's equations can then be written in phasor form as

$$\nabla \times \bar{E}_s = -j\omega\mu\bar{H}_s \quad (2.7)$$

$$\nabla \times \bar{H}_s = (\sigma + j\omega\varepsilon)\bar{E}_s \quad (2.8)$$

$$\nabla \cdot \bar{E}_s = 0 \quad (2.9)$$

$$\nabla \cdot \bar{H}_s = 0 \quad (2.10)$$

The phasor form of Maxwell's equations assumes that the fields are expressed in complex form as sinusoids or can be expanded in sinusoids, that is, $\bar{E} = \text{Re}\{\bar{E}_s e^{j\omega t}\}$; $\bar{H} = \text{Re}\{\bar{H}_s e^{j\omega t}\}$. Thus solutions stemming from

the use of Eqs. (2.7) to (2.10) must be sinusoidal solutions. One such solution is the Helmholtz wave equation.

2.2 The Helmholtz Wave Equation

We can solve for the propagation of waves in free space by taking the curl of both sides of Eq. (2.7) and eliminating $\vec{H}_s$ by using Eq. (2.8). The result can be re-written as

$$\nabla \times \nabla \times \vec{E}_s = -j\omega\mu(\sigma + j\omega\varepsilon)\vec{E}_s \quad (2.11)$$

We also can invoke a well-known vector identity where $\nabla \times \nabla \times \vec{E}_s = \nabla(\nabla \cdot \vec{E}_s) - \nabla^2 \vec{E}_s$. Since we are in free space, where no sources exist, the divergence of $\vec{E}_s$ equals zero as given by Eq. (2.9). Equation (2.11) can thus be rewritten as

$$\nabla^2 \vec{E}_s - \gamma^2 \vec{E}_s = 0 \quad (2.12)$$

where

$$\gamma^2 = j\omega\mu(\sigma + j\omega\varepsilon) \quad (2.13)$$

Equation (2.12) is called the *vector Helmholtz⁵ wave equation* and γ is known as the *propagation constant*. Since γ is obviously a complex quantity, it can be more simply expressed as

$$\gamma = \alpha + j\beta \quad (2.14)$$

where α is attenuation constant (Np/m) and β is phase constant (rad/m).

Through a simple manipulation of the real part of γ^2 and the magnitude of γ^2 , one can derive separate equations for α and β as given by

$$\alpha = \omega \sqrt{\frac{\mu\varepsilon}{2} \left[\sqrt{1 + \left(\frac{\sigma}{\omega\varepsilon}\right)^2} - 1 \right]} \quad (2.15)$$

$$\beta = \omega \sqrt{\frac{\mu\varepsilon}{2} \left[\sqrt{1 + \left(\frac{\sigma}{\omega\varepsilon}\right)^2} + 1 \right]} \quad (2.16)$$

It can be seen that the attenuation constant in Eq. (2.15) and the phase constant in Eq. (2.16) are functions of the radian frequency ω , constitutive parameters μ and ε , and also of the conductivity of the medium σ . The term $\sigma/\omega\varepsilon$ is typically referred to as the *loss tangent*.

⁵Hermann Helmholtz (1821–1894): A German born physician who served in the Prussian army fighting Napoleon. He was a self-taught mathematician.

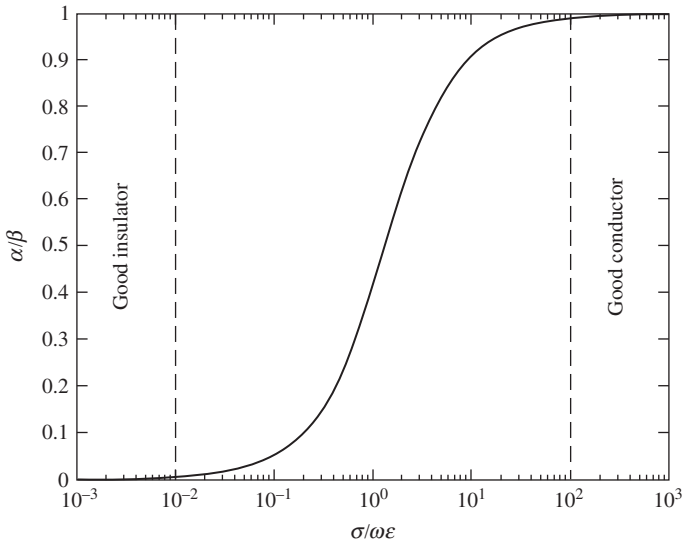


Figure 2.1 α/β vs. loss tangent.

When a medium has a loss tangent $<.01$, the material is said to be a good insulator. Indoor building materials, such as brick or concrete, have loss tangent values near .1 at 3 GHz. When the loss tangent is >100 , the material is said to be a good conductor. Figure 2.1 shows a plot of α/β vs. the loss tangent.

We can also solve for the propagation of magnetic fields in free space by taking the curl of both sides of Eq. (2.8) and substituting Eq. (2.7) to get the Helmholtz equation for $\bar{H}_s$ as given by

$$\nabla^2 \bar{H}_s - \gamma^2 \bar{H}_s = 0 \tag{2.17}$$

The propagation constant is identical to that given in Eq. (2.14).

2.3 Propagation in Rectangular Coordinates

The vector Helmholtz equation in Eq. (2.12) can be solved in any orthogonal coordinate system by substituting the appropriate del (∇) operator for that coordinate system. Let us first assume a solution in rectangular coordinates. Figure 2.2 shows a rectangular coordinate system relative to the earth's surface.

It is assumed that the z axis is perpendicular to the surface whereas the x and y coordinates are parallel. Let us also assume that the electric field is polarized in the z -direction and is only propagating in the

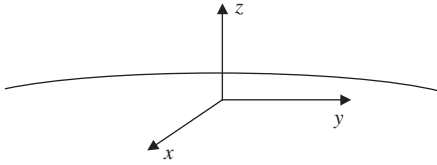


Figure 2.2 Rectangular coordinate system relative to the earth.

x -direction. Thus Eq. (2.12) can be furthered simplified.

$$\frac{d^2 E_{xs}}{dx^2} - \gamma^2 E_{xs} = 0 \quad (2.18)$$

The solution is of the form

$$E_{zs}(x) = E_0 e^{-\gamma x} + E_1 e^{\gamma x} \quad (2.19)$$

Assuming that the field propagates only in the positive x -direction and is finite at infinity then E_1 must be equal to 0 giving

$$E_{zs}(x) = E_0 e^{-\gamma x} \quad (2.20)$$

We can revert the phasor of Eq. (2.20) back to the time domain by reintroducing $e^{j\omega t}$. Thus

$$\bar{E}(x, t) = \text{Re}\{E_0 e^{-\gamma x} e^{j\omega t} \hat{z}\} = \text{Re}\{E_0 e^{-\alpha x} e^{j(\omega t - \beta x)} \hat{z}\}$$

or

$$\bar{E}(x, t) = E_0 e^{-\alpha x} \cos(\omega t - \beta x) \hat{z} \quad (2.21)$$

Figure 2.3 shows an example plot of the normalized propagating E -field at a fixed point in time.

The attenuation constant in Eq. (2.21) is representative of a medium with an ideal homogeneous conductivity. In a more realistic radio wave

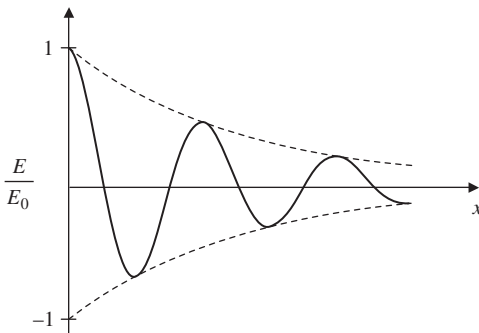


Figure 2.3 Propagating E -field.

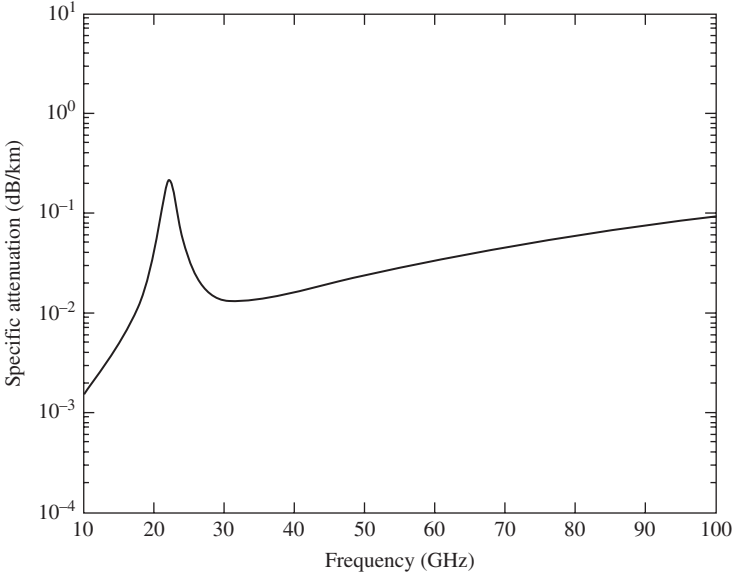


Figure 2.4 Attenuation by water vapor at sea level.

propagation model, the attenuation is further affected by the presence of atmospheric gases, clouds, fog, rain, and water vapor. This is especially true for propagation frequencies above 10 GHz. Thus a more sophisticated model is necessary to more accurately represent the full signal attenuation. Three good references explaining atmospheric attenuation, due to other factors, are Collin [4], Ulaby, Moore, Fung [5], and Elachi [6]. Figure 2.4 shows the attenuation by the molecular resonance of uncondensed water vapor. A resonance condition can be seen to occur at about 22 GHz. The equation for plotting Fig. 2.4 was taken from Frey [7].

2.4 Propagation in Spherical Coordinates

We may also calculate the propagation of electric fields from an isotropic point source in spherical coordinates. The traditional wave equation approach is developed through vector and scalar potentials as shown in Collin [4] or Balanis [8]. However, a thumbnail derivation, albeit less rigorous, can be taken directly from Eq. (2.12) in spherical coordinates.⁶ This derivation assumes an isotropic point source. Figure 2.5 shows a spherical coordinate system over the earth.

⁶In the far-field $\vec{E} = -j\omega\vec{A}$, thus the Helmholtz wave equation is of the same form for either $\vec{E}$ or $\vec{A}$.

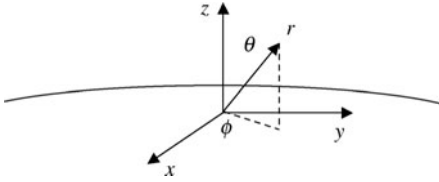


Figure 2.5 Spherical coordinate system relative to the earth.

We will assume that the source is isotropic such that the electric field solution is not a function of (θ, ϕ) . (It should be noted that we have made no assumption that the isotropic source is an infinitesimal dipole. This allows a simplification of the solution.) Assuming that the electric field is polarized in the θ -direction and is only a function of r , we can write Eq. (2.12) in spherical coordinates as

$$\frac{d}{dr} \left(r^2 \frac{dE_{\theta s}}{dr} \right) - \gamma^2 r^2 E_{\theta s} = 0 \quad (2.22)$$

For finite fields, the solution can be seen to be of the form

$$E_{\theta s}(r) = \frac{E_0 e^{-\gamma r}}{r} \quad (2.23)$$

As before, we can express the phasor of Eq. (2.23) in the time domain to get

$$\bar{E}(r, t) = \frac{E_0 e^{-\alpha r}}{r} \cos(\omega t - \beta r) \hat{\theta} \quad (2.24)$$

The difference between Eq. (2.23) in spherical coordinates and Eq. (2.20) in rectangular coordinates is due to the fact that there is a point source producing the propagating wave thus giving rise to the $1/r$ dependence. This factor is termed *spherical spreading* implying that since the radiation emanates from a point source, the field spreads out as if over the surface of a sphere whose radius is r . Since all finite length antennas are used to generate radio waves, all propagating electric far-fields undergo a spherical spreading loss as well as the attenuation loss due to factors discussed earlier. The solution depicted in Eqs. (2.23) and (2.24) is identical in form to the more classically derived solutions. The term E_0 can be viewed as being frequency dependent for finite length sources.

2.5 Electric Field Boundary Conditions

All electric and magnetic field behavior is influenced and disrupted by boundaries. Boundaries interrupt the normal flow of propagating fields and change the field strengths of static fields. All material

discontinuities give rise to reflected, transmitted, refracted, diffracted, and scattered fields. These perturbed fields give rise to multipath conditions to exist within a channel. As the number of material discontinuities increases, the number of multipath signals increases. Boundary conditions on electric fields must be established in order to determine the nature of the reflection, transmission, or refraction between dielectric media. Scattering or diffraction conditions are accounted for by different mechanisms. These will be discussed in Sec. 2.8.

Two of Maxwell's equation, in integral form, can be used to establish electric field boundary conditions. These are the conservation of energy as given by

$$\oint \bar{E} \cdot d\bar{\ell} = 0 \quad (2.25)$$

and the conservation of flux as given by

$$\oint \bar{D} \cdot d\bar{S} = Q_{enc} \quad (2.26)$$

Equation (2.25) can be applied to find tangential boundary conditions (E_t) and Eq. (2.26) can be applied to find normal boundary conditions (D_n). Let us recast the electric field strength and the electric flux density as having tangential and normal components relative to the boundary.

$$\bar{E} = \bar{E}_t + \bar{E}_n \quad (2.27)$$

$$\bar{D} = \bar{D}_t + \bar{D}_n \quad (2.28)$$

Figure 2.6 shows the boundary between two media and the corresponding tangential and normal electric fields on each side of the boundary.

Applying Eq. (2.25) to the loop shown in Fig. 2.6, and allowing the loop dimensions to become very small relative to the radii of curvature of the boundary, we obtain the following simplification of the line integral:

$$E_{t2}\Delta\ell - E_{n2}\frac{\Delta h}{2} - E_{n1}\frac{\Delta h}{2} - E_{t1}\Delta\ell = 0 \quad (2.29)$$

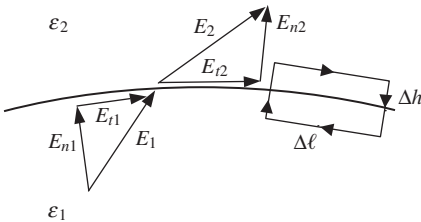


Figure 2.6 Dielectric boundary with E fields.

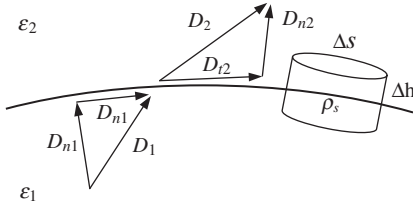


Figure 2.7 Dielectric boundary with electric flux density.

Allowing the loop height $\Delta h \rightarrow 0$, Eq. (2.29) becomes

$$E_{t1} = E_{t2} \quad (2.30)$$

Thus, tangential E is continuous across the boundary between two dielectrics.

Figure 2.7 shows the boundary between two media and the corresponding tangential and normal electric flux densities on each side of the boundary. The boundary surface has a surface charge density ρ_s .

Applying Eq. (2.26) to the cylindrical closed surface shown in Fig. 2.7, and allowing the cylinder dimensions to become very small relative to the radii of curvature of the boundary, we obtain the following simplification of the surface integral:

$$D_{n2}\Delta s - D_{n1}\Delta s = \rho_s\Delta s$$

or

$$D_{n2} - D_{n1} = \rho_s \quad (2.31)$$

Thus, normal D is discontinuous across a material boundary by the surface charge density at that point.

We can apply the two boundary conditions given in Eqs. (2.30) and (2.31) to determine the refraction properties of two dissimilar dielectric materials. Let us assume that the surface charge at the boundary between the two materials is zero ($\rho_s = 0$). Let us also construct a surface normal $\hat{n}$, pointing into region 2, as shown in Fig. 2.8. Then $\vec{E}_1$ and

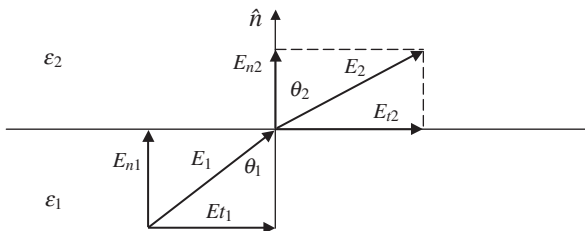


Figure 2.8 D and E at a dielectric boundary.

$\vec{D}_1$ are inclined at an angle θ_1 with respect to the surface normal. Also, $\vec{E}_2$ and $\vec{D}_2$ are inclined at an angle θ_2 with respect to the surface normal. Applying the boundary condition in Eq. (2.30) we get

$$E_1 \sin \theta_1 = E_{t1} = E_{t2} = E_2 \sin \theta_2$$

or

$$E_1 \sin \theta_1 = E_2 \sin \theta_2 \quad (2.32)$$

In the same way, we can apply a similar procedure to satisfy the boundary conditions of Eq. (2.31) to yield

$$\varepsilon_1 E_1 \cos \theta_1 = D_{n1} = D_{n2} = \varepsilon_2 E_2 \cos \theta_2$$

or

$$\varepsilon_1 E_1 \cos \theta_1 = \varepsilon_2 E_2 \cos \theta_2 \quad (2.33)$$

Dividing Eq. (2.32) by Eq. (2.33) we can perform simple algebra to obtain a relationship between the two angles of the corresponding E fields.

$$\frac{\tan \theta_1}{\tan \theta_2} = \frac{\varepsilon_{r1}}{\varepsilon_{r2}} \quad (2.34)$$

Example 2.1 Two semi-infinite dielectrics share a boundary in the $z = 0$ plane. There is no surface charge on the boundary. For $z \leq 0$, $\varepsilon_{r1} = 4$. For $z \geq 0$, $\varepsilon_{r2} = 8$.

If $\theta_1 = 30^\circ$, what is the angle θ_2 ?

Solution Let this problem be illustrated in Fig. 2.9.

Using Eq. (2.34) it can be found that

$$\theta_2 = \tan^{-1} \left(\frac{\varepsilon_{r2}}{\varepsilon_{r1}} \tan \theta_1 \right) = 49.1^\circ$$

Example 2.2 Two semi-infinite dielectrics share a boundary in the $z = 0$ plane. There is no surface charge on the boundary. For $z \leq 0$, $\varepsilon_{r1} = 4$. For

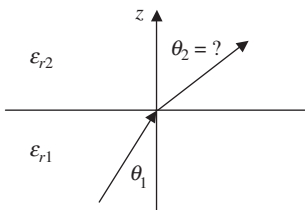


Figure 2.9 For Example 2.1.

$z \geq 0$, $\epsilon_{r2} = 8$. Given that $\vec{E}_1 = 2\hat{x} + 4\hat{y} + 6\hat{z}$, what is the electric field in region 2?

Solution We can use Eq. (2.34) by finding the angle θ_1 from the equation for $\vec{E}_1$.

However, it would be simpler to apply the boundary conditions of Eqs. (2.30) and (2.31). Thus using Eq. (2.30)

$$\vec{E}_{t1} = 2\hat{x} + 4\hat{y} = \vec{E}_{t2}$$

Also, $\vec{D}_{n2} = \epsilon_{r2}\epsilon_0\vec{E}_{n2} = 8\epsilon_0\vec{E}_{n2}\hat{z}$ and $\vec{D}_{n1} = \epsilon_{r1}\epsilon_0\vec{E}_{n1} = 4\epsilon_0(6\hat{z})$. Thus using Eq. (2.31) $E_{n2} = 3$ yielding

$$\vec{E}_2 = 2\hat{x} + 4\hat{y} + 3\hat{z}$$

2.6 Magnetic Field Boundary Conditions

The magnetic field boundary conditions are duals of the boundary conditions for electric fields. The remaining two Maxwell's equations, in integral form, can be used to establish these magnetic boundary conditions. These are Ampere's circuital law as given by

$$\oint \vec{H} \cdot d\vec{\ell} = I \quad (2.35)$$

and the conservation of magnetic flux as given by

$$\oint \vec{B} \cdot d\vec{S} = 0 \quad (2.36)$$

Equation (2.35) can be applied to find tangential boundary conditions (H_t) and Eq. (2.36) can be applied to find normal boundary conditions (B_n). Let us recast the magnetic field intensity and the magnetic flux density as having tangential and normal components relative to the magnetic boundary.

$$\vec{H} = \vec{H}_t + \vec{H}_n \quad (2.37)$$

$$\vec{B} = \vec{B}_t + \vec{B}_n \quad (2.38)$$

Figure 2.10 shows the boundary between two media and the corresponding tangential and normal magnetic fields on each side of the boundary. In addition, a surface current density $\vec{K}$ flows along the boundary.

Applying Eq. (2.37) to the loop shown in Fig. 2.10, and allowing the loop dimensions to become very small relative to the radii of curvature of the boundary, we obtain the following simplification of the line integral:

$$H_{t2}\Delta\ell - H_{n2}\frac{\Delta h}{2} - H_{n1}\frac{\Delta h}{2} - H_{t1}\Delta\ell = K\Delta\ell \quad (2.39)$$

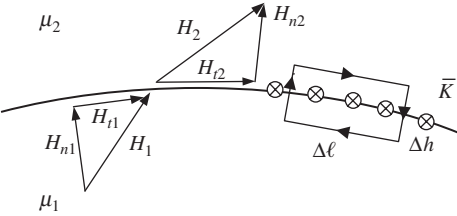


Figure 2.10 Magnetic boundary with E fields.

Allowing the loop height $\Delta h \rightarrow 0$, Eq. (2.39) becomes

$$H_{t1} - H_{t2} = K \tag{2.40}$$

Thus, tangential H is discontinuous across the boundary between two magnetic materials.

Figure 2.11 shows the boundary between two media and the corresponding tangential and normal magnetic flux densities on each side of the boundary. The boundary has no corresponding magnetic surface charge because magnetic monopoles do not exist.

Applying Eq. (2.36) to the cylindrical closed surface shown in Fig. 2.11, and allowing the cylinder dimensions to become very small relative to the radii of curvature of the boundary, we obtain the following simplification of the surface integral:

$$B_{n2} \Delta s - B_{n1} \Delta s = 0$$

or

$$B_{n2} = B_{n1} \tag{2.41}$$

Thus, normal B is continuous across a magnetic material boundary.

We can apply the two boundary conditions given in Eq. (2.41) to determine the magnetic refractive properties of two dissimilar magnetic materials. Let us assume that the surface current density at the boundary between the two materials is zero ($K = 0$). Performing a similar

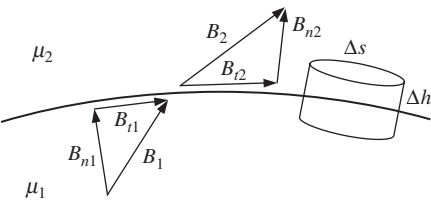


Figure 2.11 Dielectric boundary with electric flux density.

operation as in Sec. 2.5, we can use simple algebra to obtain a relationship between the two angles of the corresponding flux lines such that

$$\frac{\tan \theta_1}{\tan \theta_2} = \frac{\mu_{r1}}{\mu_{r2}} \quad (2.42)$$

Examples 2.1 and 2.2 can be applied to magnetic fields with the same corresponding values for the relative permeabilities achieving the same results. This is not surprising since electric and magnetic fields and electric and magnetic media are duals of each other.

2.7 Planewave Reflection and Transmission Coefficients

Multipath signals are the consequence of the transmitted signal reflecting, transmitting, and diffracting from various structures along the route to the receiver. In this section we will deal exclusively with the reflection and transmission of planewaves. One aspect of calculating each multipath term is being able to predict the reflection and transmission through various materials. The boundary conditions of Eqs. (2.30) and (2.40) can be invoked to allow us to determine the reflection and transmission coefficients. The simplest case is to predict the reflection and transmission across a planar boundary at normal incidence. The details of the derivation can be found in Sadiku [1].

2.7.1 Normal incidence

Figure 2.12 shows a plane wave normally incident upon a planar material boundary.

E_{is} and H_{is} symbolize the incident fields, in phasor form, propagating in the positive z -direction. E_{rs} and H_{rs} symbolize the reflected fields propagating in the minus z -direction. E_{ts} and H_{ts} symbolize the transmitted fields propagating in the positive z -direction. The exact expressions for the E and H fields are given by the following expressions:

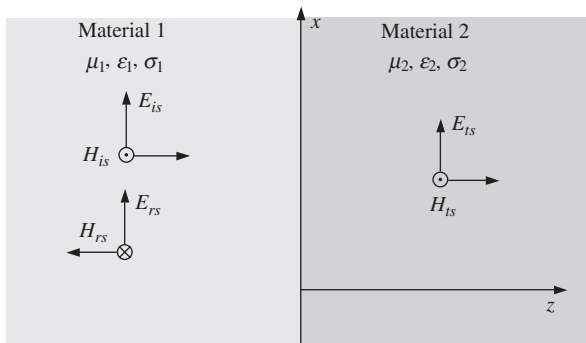


Figure 2.12 Plane wave normally incident on a material boundary at $z = 0$.

Incident fields:

$$\bar{E}_{is}(z) = E_{i0}e^{-\gamma_1 z} \hat{x} \quad (2.43)$$

$$\bar{H}_{is}(z) = \frac{E_{i0}}{\eta_1} e^{-\gamma_1 z} \hat{y} \quad (2.44)$$

Reflected fields:

$$\bar{E}_{rs}(z) = E_{r0}e^{\gamma_1 z} \hat{x} \quad (2.45)$$

$$\bar{H}_{rs}(z) = -\frac{E_{r0}}{\eta_1} e^{\gamma_1 z} \hat{y} \quad (2.46)$$

Transmitted fields:

$$\bar{E}_{ts}(z) = E_{t0}e^{-\gamma_2 z} \hat{x} \quad (2.47)$$

$$\bar{H}_{ts}(z) = \frac{E_{t0}}{\eta_2} e^{-\gamma_2 z} \hat{y} \quad (2.48)$$

where the intrinsic impedances are given by

$$\eta_1 = \sqrt{\frac{\frac{\mu_1}{\epsilon_1}}{1 - j\frac{\sigma_1}{\omega\epsilon_1}}} = \text{intrinsic impedance of medium 1}$$

$$\eta_2 = \sqrt{\frac{\frac{\mu_2}{\epsilon_2}}{1 - j\frac{\sigma_2}{\omega\epsilon_2}}} = \text{intrinsic impedance of medium 2}$$

Obviously the intrinsic impedances are dependent upon the loss tangent as well as the propagation constants in both media.

Assuming that there is no surface current at the boundary and utilizing the tangential boundary conditions of Eqs. (2.30) and (2.40), one can derive the reflection and transmission coefficients, respectively, as

$$R = \frac{\eta_2 - \eta_1}{\eta_2 + \eta_1} = |R|e^{j\theta_R} \quad (2.49)$$

$$T = \frac{2\eta_2}{\eta_2 + \eta_1} = |T|e^{j\theta_T} \quad (2.50)$$

Knowing the reflection and transmission coefficients R and T , one can determine the total electric field in regions 1 and 2. The total electric field in region 1 is given by

$$\bar{E}_{1s} = \bar{E}_{is} + \bar{E}_{rs} = E_{i0}[e^{-\gamma_1 z} + \text{Re}^{\gamma_1 z}] \hat{x} \quad (2.51)$$

whereas the total electric field in region 2 is given by

$$\bar{E}_{2s} = TE_0e^{-\gamma_2 z} \hat{x} \quad (2.52)$$

When a nonzero reflection coefficient exists and region 1 is lossless, a standing wave is established. This standing wave gives rise to an interference pattern, which is a function of distance from the boundary. This interference is a trivial example of fading that occurs in many wireless applications. It is instructive to derive this standing wave envelope.

The total field in region 1 can be reexpressed using the polar form of the reflection coefficient. If we assume that region 1 is lossless (i.e., $\sigma_1 = 0$). Then

$$\bar{E}_{1s} = E_{i0} [e^{-j\beta_1 z} + |R|e^{j(\beta_1 z + \theta_R)}] \hat{x} \quad (2.53)$$

Combining the real and imaginary parts in Eq. (2.53) we get

$$\begin{aligned} \bar{E}_{1s} = E_{i0} [& \cos(\beta_1 z) + |R| \cos(\beta_1 z + \theta_R) \\ & + (|R| \sin(\beta_1 z + \theta_R) - \sin(\beta_1 z))e^{j(\pi/2)}] \hat{x} \end{aligned} \quad (2.54)$$

We may now convert the phasor of Eq. (2.54) into instantaneous time form

$$\begin{aligned} E_1(z, t) = E_{i0} [& (\cos(\beta_1 z) + |R| \cos(\beta_1 z + \theta_R)) \cos \omega t \\ & - (|R| \sin(\beta_1 z + \theta_R) - \sin(\beta_1 z)) \sin \omega t] \end{aligned} \quad (2.55)$$

Since Eq. (2.55) contains two components that are in phase quadrature, we may easily find the magnitude to be given as

$$\begin{aligned} |E_1(z)| &= E_{i0} \sqrt{(\cos(\beta_1 z) + |R| \cos(\beta_1 z + \theta_R))^2 + (\sin(\beta_1 z) - |R| \sin(\beta_1 z + \theta_R))^2} \\ &= E_{i0} \sqrt{1 + |R|^2 + 2|R| \cos(2\beta_1 z + \theta_R)} \end{aligned} \quad (2.56)$$

Equation (2.56) has extrema when the cosine term is either $+1$ or -1 . Thus the maximum and minimum values are given as

$$|E_1|_{\max} = E_{i0} \sqrt{1 + |R|^2 + 2|R|} = E_{i0}(1 + |R|) \quad (2.57)$$

$$|E_1|_{\min} = E_{i0} \sqrt{1 + |R|^2 - 2|R|} = E_{i0}(1 - |R|) \quad (2.58)$$

The standing wave ratio s is defined as the ratio of $|E_1|_{\max}/|E_2|_{\min}$.

Example 2.3 A boundary exists between two regions where region 1 is free space and region 2 has the parameters $\mu_2 = \mu_0$, $\varepsilon_2 = 4\varepsilon_0$, and $\sigma_2 = 0$. If $E_{i0} = 1$, use MATLAB to plot the standing wave pattern over the range $-4\pi < \beta_1 z < 0$.

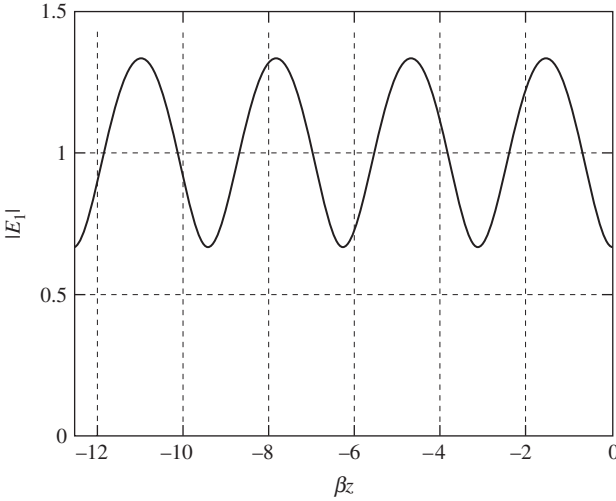


Figure 2.13 Standing wave pattern for normal incidence ($\eta_2 = \eta_1/2$).

Solution Solving for the reflection coefficient

$$R = \frac{\sqrt{\frac{\mu_0}{4\epsilon_0}} - \sqrt{\frac{\mu_0}{\epsilon_0}}}{\sqrt{\frac{\mu_0}{4\epsilon_0}} + \sqrt{\frac{\mu_0}{\epsilon_0}}} = -\frac{1}{3} = \frac{1}{3}e^{j\pi}$$

Using Eq. (2.56) and MATLAB, the standing wave pattern appears as shown in Fig. 2.13.

Normal incidence is a special case of the more interesting oblique incidence. Oblique incidence will be discussed in the next section.

2.7.2 Oblique incidence

The oblique incidence case is considerably more complicated than the normal incidence case and an intensive derivation of the reflection and transmission coefficients can be seen in Sadiku [1]. The oblique incidence reflection and transmission coefficients are called the *Fresnel coefficients*. Only the highlights are given in this discussion.

Figure 2.14 depicts an incident field upon a boundary. It is assumed that both media are lossless. The electric field is parallel to the plane of incidence. The plane of incidence is that plane containing the surface normal and the direction of propagation. The angles θ_i , θ_r , and θ_t are the angles of incidence, reflection, and transmission with respect to the surface normal ($\pm x$ axis). By a careful application of the boundary conditions, given in Eqs. (2.30) and (2.40), one can determine two laws.

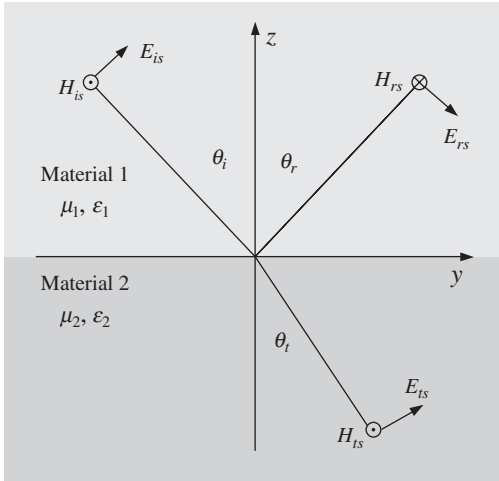


Figure 2.14 Parallel polarization reflection and transmission.

The first is Snell's law of reflection, which states that the angle of reflection equals the angle of incidence. (This property is also termed *specular* reflection.)

$$\theta_r = \theta_i \quad (2.59)$$

The second result is the consequence of the conservation of phase and is also called *Snell's law of refraction*.

$$\beta_1 \sin \theta_i = \beta_2 \sin \theta_t \quad (2.60)$$

Parallel polarization. The incident field, for the parallel polarization case, is indicated in Fig. 2.14. The coordinate system is rotated from the coordinate system in Fig. 2.12 so as to indicate reflection from a horizontal surface. This is often the case for elevated antennas over a flat earth. This is a parallel-polarized field because the E field is in the y - z plane, which is the plane of incidence.

The incident, reflected, and transmitted electric fields are given by

$$\vec{E}_{is} = E_0(\cos \theta_i \hat{y} + \sin \theta_i \hat{z})e^{-j\beta_1(y \sin \theta_i - z \cos \theta_i)} \quad (2.61)$$

$$\vec{E}_{rs} = R_{||} E_0(\cos \theta_i \hat{y} - \sin \theta_i \hat{z})e^{-j\beta_1(y \sin \theta_i + z \cos \theta_i)} \quad (2.62)$$

$$\vec{E}_{ts} = T_{||} E_0(\cos \theta_t \hat{y} + \sin \theta_t \hat{z})e^{-j\beta_2(y \sin \theta_t - z \cos \theta_t)} \quad (2.63)$$

where the reflection and transmission coefficients are given as

$$R_{||} = \frac{\eta_2 \cos \theta_t - \eta_1 \cos \theta_i}{\eta_2 \cos \theta_t + \eta_1 \cos \theta_i} \quad (2.64)$$

and

$$T_{||} = \frac{2\eta_2 \cos \theta_i}{\eta_2 \cos \theta_t + \eta_1 \cos \theta_i} \tag{2.65}$$

The term $\cos \theta_t$ in Eqs. (2.64) and (2.65) can be easily calculated using Eq. (2.60) to be

$$\cos \theta_t = \sqrt{1 - \sin^2 \theta_t} = \sqrt{1 - \frac{\mu_1 \varepsilon_1}{\mu_2 \varepsilon_2} \sin^2 \theta_i} \tag{2.66}$$

Figure 2.15 shows a plot of the magnitude of the reflection and transmission coefficients in the case that both media are nonmagnetic and lossless. The permittivities are given respectively as $\varepsilon_1 = \varepsilon_0$ and $\varepsilon_2 = 2\varepsilon_0, 8\varepsilon_0, 32\varepsilon_0$.

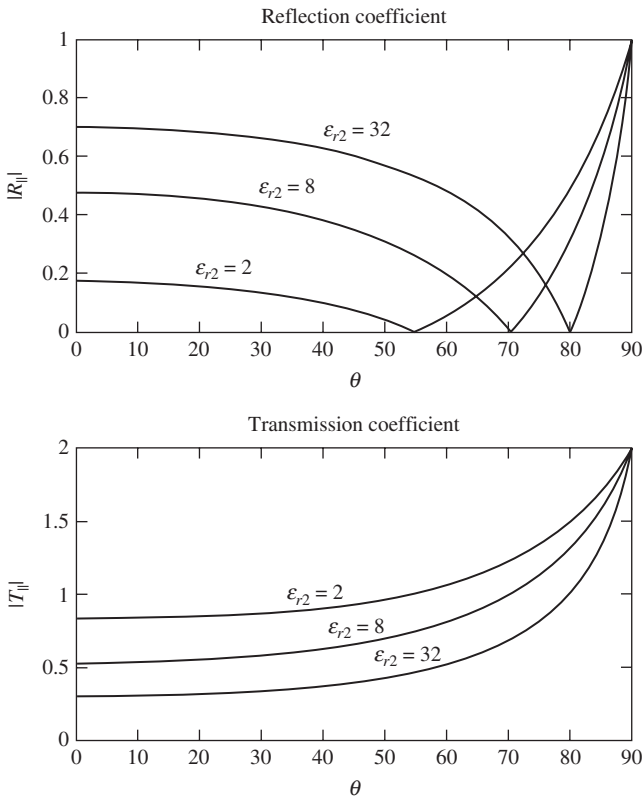


Figure 2.15 Reflection and transmission coefficient magnitude for parallel polarization.

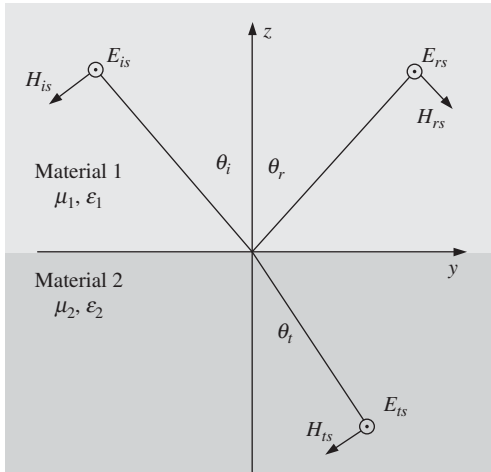


Figure 2.16 Perpendicular polarization reflection and transmission.

Perpendicular polarization. The incident field, for the parallel polarization case, is indicated in Fig. 2.16.

The incident, reflected, and transmitted electric fields are given by

$$\bar{E}_{is} = E_0 \hat{x} e^{-j\beta_1(y \sin \theta_i - z \cos \theta_i)} \quad (2.67)$$

$$\bar{E}_{rs} = R_{\perp} E_0 \hat{x} e^{-j\beta_1(y \sin \theta_i + z \cos \theta_i)} \quad (2.68)$$

$$\bar{E}_{ts} = T_{\perp} E_0 \hat{x} e^{-j\beta_2(y \sin \theta_t - z \cos \theta_t)} \quad (2.69)$$

where the reflection and transmission coefficients are given as

$$R_{\perp} = \frac{\eta_2 \cos \theta_i - \eta_1 \cos \theta_t}{\eta_2 \cos \theta_i + \eta_1 \cos \theta_t} \quad (2.70)$$

and

$$T_{\perp} = \frac{2\eta_2 \cos \theta_i}{\eta_2 \cos \theta_i + \eta_1 \cos \theta_t} \quad (2.71)$$

Figure 2.17 shows a plot of the magnitude of the reflection and transmission coefficients in the case that both media are nonmagnetic and lossless. The permittivities are given as $\epsilon_1 = \epsilon_0$ and $\epsilon_2 = 2\epsilon_0, 8\epsilon_0, 32\epsilon_0$, respectively.

2.8 Propagation Over Flat Earth

Having discussed the planewave reflection coefficients for parallel and perpendicular polarization, we are now in a position to analyze the propagation of planewaves over flat earth. Even though the earth has

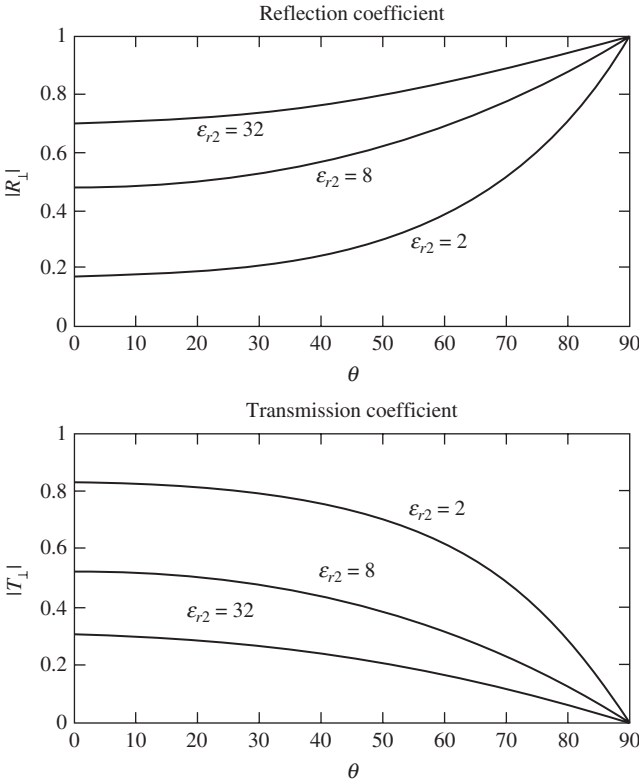


Figure 2.17 Reflection and transmission coefficient magnitude for perpendicular polarization.

curvature, and this curvature dramatically affects long distance propagation, we will limit this discussion to short distances and assume that the earth is flat. This allows us to make some propagation generalizations. This topic is a critical start in understanding general multipath propagation problems because the flat earth model allows us to include a second indirect path. This second path will produce interference effects at the receiver.

Let us consider isotropic transmit and receive antennas as shown in Fig. 2.18. The transmitting antenna is at height h_1 , and the receiving antenna is at height h_2 , and the two antennas are separated by a horizontal distance d . The reflection coefficient from the surface of the earth is R and can be approximated either by Eq. (2.64) or (2.70). It should be noted that R is complex for most earth reflections.

The received signal at the receiver is composed of the direct path propagation as well as the reflected signal at the point y . Thus, the

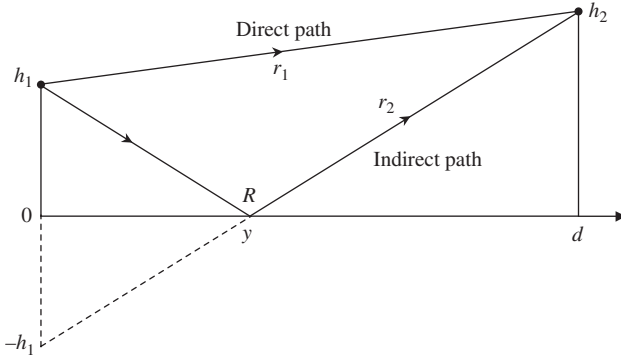


Figure 2.18 Flat earth model with two isotropic antennas.

composite signal, due to the direct and indirect paths is proportional to the following equation:

$$\frac{e^{-jkr_1}}{r_1} + R \frac{e^{-jkr_2}}{r_2} \quad (2.72)$$

where k is the wavenumber as given by the dispersion relation

$$k^2 = k_x^2 + k_y^2 + k_z^2 = \beta^2 \quad (2.73)$$

with $k = \omega\sqrt{\mu\epsilon} = 2\pi/\lambda$

The reflection coefficient R is normally complex and can alternatively be expressed as $R = |R|e^{-j\psi}$. Through some simple algebra, it can be shown that

$$r_1 = \sqrt{d^2 + (h_2 - h_1)^2} \quad (2.74)$$

$$r_2 = \sqrt{d^2 + (h_2 + h_1)^2} \quad (2.75)$$

Factoring the direct path term out of Eq. (2.72), we get

$$\frac{e^{-jkr_1}}{r_1} \left[1 + R \frac{r_1}{r_2} e^{-jk(r_2 - r_1)} \right] \quad (2.76)$$

The magnitude of the second term in Eq. (2.76) is called the *path gain factor* F . F is also similar to the *height gain* defined in Bertoni [9]. Thus

$$F = \left| 1 + R \frac{r_1}{r_2} e^{-jk(r_2 - r_1)} \right| \quad (2.77)$$

This factor is analogous to the array factor for a two-element array separated by a distance of $2h_1$. In addition, the point of reflection at y is the solution to a simple algebra problem and is given as

$y = dh_1/(h_1 + h_2)$. Using the definition of the path gain factor, we can rewrite Eq. (2.76) as

$$\frac{e^{-jkr_1}}{r_1} F \quad (2.78)$$

If we make the assumption that the antenna heights $h_1, h_2 \ll r_1, r_2$, we can use a binomial expansion on Eqs. (2.74) and (2.75) to simplify the distances.

$$r_1 = \sqrt{d^2 + (h_2 - h_1)^2} \approx d + \frac{(h_2 - h_1)^2}{2d} \quad (2.79)$$

$$r_2 = \sqrt{d^2 + (h_2 + h_1)^2} \approx d + \frac{(h_2 + h_1)^2}{2d} \quad (2.80)$$

Consequently, the difference in path lengths is given as

$$r_2 - r_1 = \frac{2h_1h_2}{d} \quad (2.81)$$

We additionally also can assume that $r_1/r_2 \approx 1$ for long distances. Under these conditions, we have a shallow grazing angle at the point of reflection. Thus, $R \approx -1$. By substituting R and Eq. (2.81) into the path gain factor, we now can get

$$F = 2 \left| \sin \frac{kh_1h_2}{d} \right| = 2 \left| \sin \frac{2\pi h_1}{d} \frac{h_2}{\lambda} \right| \quad (2.82)$$

Figure 2.19 shows a typical plot of F for a range of values of h_2/λ where $h_1 = 5$ m, $d = 200$ m. It is clear that the received signal can vary between

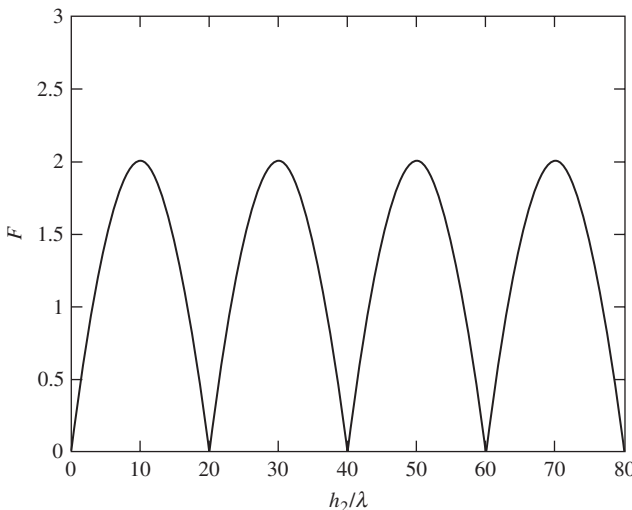


Figure 2.19 Path gain factor.

zero and twice the direct path signal strength due to constructive and destructive interference created by the indirect path. The path gain factor can also be used to create a coverage diagram. The details of which can be found in [4].

2.9 Knife-Edge Diffraction

In addition to received fields being disrupted by reflection from the ground, additional propagation paths can be created by diffraction from hills, buildings, and other objects. These structures may not be positioned at angles so as to allow specular reflection; however, they might provide for a diffraction term in the total received field. Figure 2.20 shows a hill of height h located between the transmit and receive antennas. This hill can be modeled as a half-plane or knife edge. It is assumed that there is no specular reflection from the hilltop. It is also assumed that the hill blocks any possible reflections from the ground, which arrive at the receiving antenna. Thus, the received field is only composed of the direct path and the diffraction path terms. h_c is the clearance height from the knife-edge to the direct path. $h_c < 0$ corresponds to the knife-edge being below the line of sight and therefore two propagation paths exist as shown in the figure. When $h_c > 0$ the knife-edge obstructs the direct path. Thus only the diffraction term is received. d_1 and d_2 are the respective horizontal distances to the knife-edge plane ($d = d_1 + d_2$). The diffracted field can allow for a received signal even when line of sight is not possible. If the receive

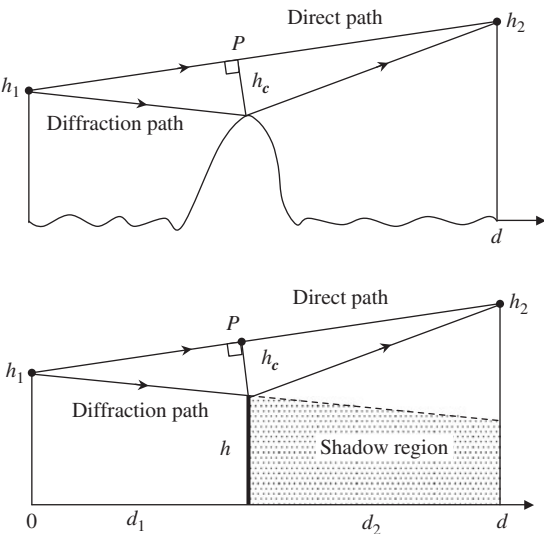


Figure 2.20 Knife-edge diffraction from a hill.

antenna has no direct path to the transmit antenna, the receive antenna is said to be in the shadow region. If the receiver is in the line of sight of the transmitter, it is said to be in the lit region. The derivation of the solution for the diffracted field can be found in Collin [4] or in Jordan and Balmain [10]. It can be shown that the path gain factor due to diffraction is given by

$$F_d = \frac{1}{\sqrt{2}} \left| \int_{-H_c}^{\infty} e^{-j\pi u^2/2} du \right| \tag{2.83}$$

where

$$H_c \approx h_c \sqrt{\frac{2d}{\lambda d_1 d_2}}$$

Thus, we can replace the path gain factor F in the flat-earth model with the path gain factor for diffraction. We can therefore rewrite Eq. (2.78) using F_d to become

$$\frac{e^{-jkr}}{r} F_d \tag{2.84}$$

Figure 2.21 shows the path gain factor F_d for a range of H_c values. It can be seen that when the knife-edge is below the line of sight ($h_c < 0$), there is an interference pattern as the direct path and the diffracted path phase in and out. However, when the knife-edge is above the line

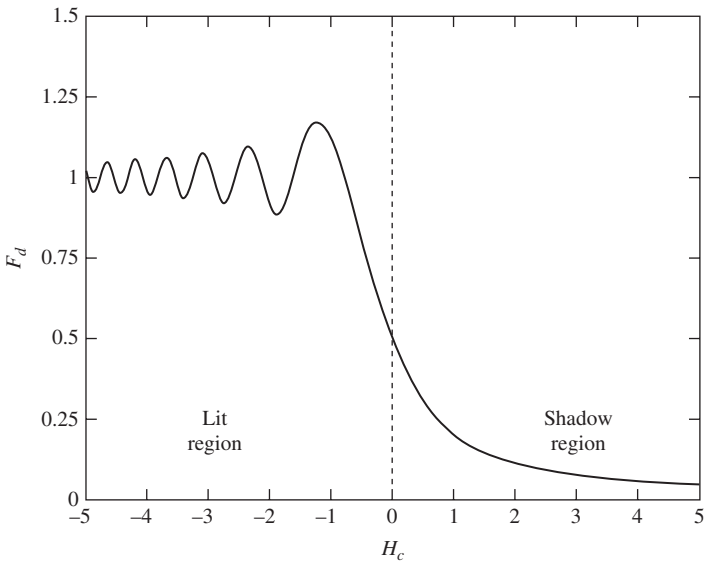


Figure 2.21 Path gain factor for knife-edge diffraction.

of sight ($h_c > 0$) no direct path exists and the field quickly diminishes deeper into the shadow region. In the case where $h_c = 0$, the path gain factor is .5 showing that the received field is 6 dB down from the direct path alone.

References

1. Sadiku, M.N.O., *Elements of Electromagnetics*, 3d ed., Oxford University Press, Oxford, 2001.
2. Hayt, W.H., Jr., and J.A. Buck, *Engineering Electromagnetics*, 6th ed., McGraw-Hill, New York, 2001.
3. Ulaby, F.T., *Fundamentals of Applied Electromagnetics*, Media ed., Prentice Hall, New York, 2004.
4. Collin, R.E., *Antennas and Radiowave Propagation*, McGraw-Hill, New York, 1985.
5. Ulaby, F.T., R.K. Moore, and A.K. Fung, *Microwave Remote Sensing Fundamentals and Radiometry*, Vol. I, Artech House, Boston, MA, 1981.
6. Elachi, C., *Introduction to the Physics and Techniques of Remote Sensing*, Wiley Interscience, New York, 1987.
7. Frey, T.L., Jr., "The Effects of the Atmosphere and Weather on the Performance of a mm-Wave Communication Link," *Applied Microwave & Wireless*, Vol. 11, No. 2, pp. 76–80, Feb. 1999.
8. Balanis, C., *Antenna Theory Analysis and Design*, 2d ed., Wiley, New York, 1997.
9. Bertoni, H., *Radio Propagation for Modern Wireless Systems*, Prentice Hall, New York, 2000.
10. Jordan, E., and K. Balmain, *Electromagnetic Waves and Radiating Systems*, 2d ed., Prentice Hall, New York, 1968.

Problems

- 2.1** For a lossy medium ($\sigma \neq 0$) with the following constitutive parameters, $\mu = 4\mu_0$, $\varepsilon = 2\varepsilon_0$, $\sigma/\omega\varepsilon = 1$, $f = 1$ MHz, find α and β .
- 2.2** For a lossy material such that $\mu = 6\mu_0$, $\varepsilon = \varepsilon_0$. If the attenuation constant is 1 Np/m at 10 MHz, find
- (a) The phase constant β
 - (b) The loss tangent
 - (c) The conductivity σ
 - (d) The intrinsic impedance
- 2.3** Use MATLAB to plot the ratio α/β for the range of $.01 < \sigma/\omega\varepsilon < 100$ with the horizontal scale being log base 10.
- 2.4** Using Eq. (2.21), if $\mu = \mu_0$, $\varepsilon = 4\varepsilon_0$, $\sigma/\omega\varepsilon = 1$, $f = 100$ MHz, how far must the wave travel in the z direction before the amplitude is attenuated by 30 percent?
- 2.5** Two semi-infinite dielectrics share a boundary in the $z = 0$ plane as depicted in Fig. P2.1. There is no surface charge on the boundary. For $z \leq 0$, $\varepsilon_{r1} = 2$. For $z \geq 0$, $\varepsilon_{r2} = 6$. If $\theta_1 = 45^\circ$, what is the angle θ_2 ?

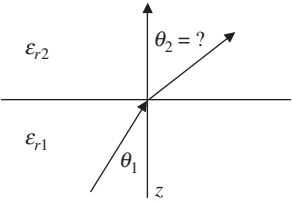


Figure P2.1 For Prob. 2.5.

2.6 Two semi-infinite dielectrics share a boundary in the $z = 0$ plane. There is no surface charge on the boundary. For $z \leq 0$, $\epsilon_{r1} = 2$. For $z \geq 0$, $\epsilon_{r2} = 4$. Given that $\vec{E}_1 = 4\hat{x} + 2\hat{y} + 3\hat{z}$ in region 1, find the electric field in region 2 ($\vec{E}_2$)?

2.7 Two semi-infinite magnetic regions share a boundary in the $z = 0$ plane. There is no surface current density K on the boundary. For $z \leq 0$, $\mu_{r1} = 4$. For $z \geq 0$, $\mu_{r2} = 2$. Given that $\vec{H}_1 = 4\hat{x} + 2\hat{y} + 3\hat{z}$ in region 1, what is the magnetic flux density in region 2 ($\vec{B}_2$)?

2.8 A plane wave normally incident on a material boundary as indicated in Fig. P2.2. Both regions are nonmagnetic. $\epsilon_{r1} = 2$, $\epsilon_{r2} = 4$. If the loss tangent in region 1 is 1.732 and the loss tangent in region 2 is 2.8284, find the following:

- (a) The intrinsic impedance in region 1
- (b) The intrinsic impedance in region 2
- (c) The reflection coefficient R
- (d) The transmission coefficient T

2.9 Modify Prob. 2.8 such that the loss tangent in region 1 is zero. If $E_{i0} = 1$, use MATLAB to plot the standing wave pattern for the range $-2\pi < \beta z < 0$.

2.10 In the oblique incidence case with parallel polarization, as depicted in Fig. 2.14, both regions are nonmagnetic and $\epsilon_{r1} = 1$, $\epsilon_{r2} = 6$. If the loss tangent

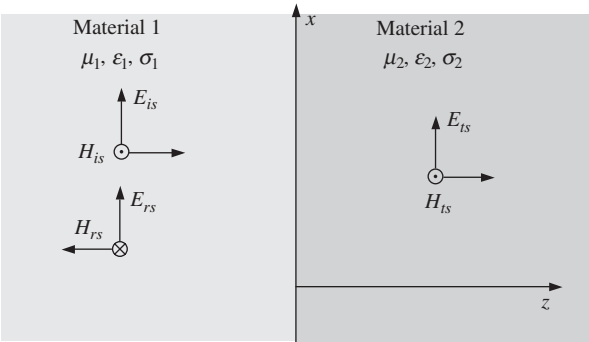


Figure P2.2 Plane wave at normal incidence on a boundary.

in region 1 is 1.732 and the loss tangent in region 2 is 2.8284, $\theta_i = 45^\circ$. Find the following:

- The intrinsic impedance in region 1
- The intrinsic impedance in region 2
- The reflection coefficient $R_{||}$
- The transmission coefficient $T_{||}$

2.11 Repeat Prob. 2.10 but for the perpendicular polarization case as depicted in Fig. 2.16.

2.12 Region 1 is free space and region 2 is lossless and nonmagnetic with $\epsilon_{r2} = 2, 8, 64$. Use MATLAB and superimpose plots for all three dielectric constants as indicated below for a range of angles $0 < \theta < 90^\circ$

- $R_{||}$
- $T_{||}$
- $R_{\perp}$
- $T_{\perp}$

2.13 For two antennas over flat earth, derive the equation for the specular reflection point y in terms of d , h_1 , and h_2 .

2.14 Using Fig. 2.18 with $h_1 = 20$ m, $h_2 = 200$ m, and $d = 1$ km, $R = -1$, use MATLAB to plot the exact path gain factor [Eq. (2.77)] for 200 MHz $< f < 400$ MHz. Assume free space.

2.15 Repeat Prob. 2.14 but allow $f = 300$ MHz and plot F for the range 500 m $< d < 1000$ m. Assume free space.

2.16 Repeat Prob. 2.14 but allow the incident field be perpendicular to the plane of incidence such that E is horizontal to the ground. Allow the ground to be nonmagnetic and have a dielectric constant $\epsilon_r = 4$.

- What is the Fresnel reflection coefficient?
- Plot the path gain factor F for 200 MHz $< f < 400$ MHz.

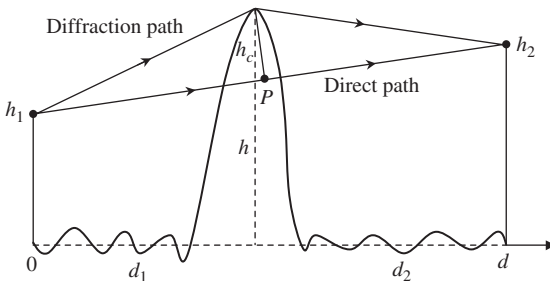


Figure P2.3 Hill between two antennas.

2.17 Use MATLAB and plot the path gain factor due to diffraction (F_d) found in Eq. (2.82) for the range of values $-8 < H_c < 8$.

2.18 For the hill between two antennas as shown in Fig. P2.3 where $h_1 = 150$ m, $h_2 = 200$ m, $d_1 = 300$ m, $d_2 = 700$ m, $d = 1$ km, $h = 250$ m, $f = 300$ MHz.

- (a) What is the direct path length r_1 ?
- (b) What is the indirect path length r_2 ?
- (c) What is the clearance height h_c (be sure your sign is right)?
- (d) What is the integration height H_c ?
- (e) Using Eq. (2.84), what is the magnitude of the received signal if $E_0 = 1$ V/m?

Antenna Fundamentals

The design and analysis of smart antennas assumes a working knowledge of many different but related disciplines. A smart antenna designer must rely on such disciplines as (1) random processes, (2) electromagnetics, (3) propagation, (4) spectral estimation methods, (5) adaptive techniques, and (6) antenna fundamentals. Especially though, smart antenna design is heavily dependent upon a basic knowledge of antenna theory. It is critical to match the individual antenna behavior with the overall system requirements. Thus, this chapter will cover the relevant antenna topics such as the near and far fields that surround antennas, power densities, radiation intensities, directivities, beamwidths, antenna reception, and fundamental antenna designs including dipoles and loops. It is not necessarily required that one have an extensive antenna background, in order to understand smart antennas, but it would be wise to review some of the literature referenced in this chapter. The foundation for much of the material in this chapter is taken from the books of Balanis [1], Kraus and Marhefka [2], and Stutzman and Thiele [3].

3.1 Antenna Field Regions

Antennas produce complex electromagnetic (EM) fields both near to and far from the antennas. Not all of the EM fields generated actually radiate into space. Some of the fields remain in the vicinity of the antenna and are viewed as reactive near fields; much the same way as an inductor or capacitor is a reactive storage element in lumped element circuits. Other fields do radiate and can be detected at great distances. An excellent treatment of the antenna field regions is given in [1] and [2]. Figure 3.1 shows a simple dipole antenna with four antenna regions.

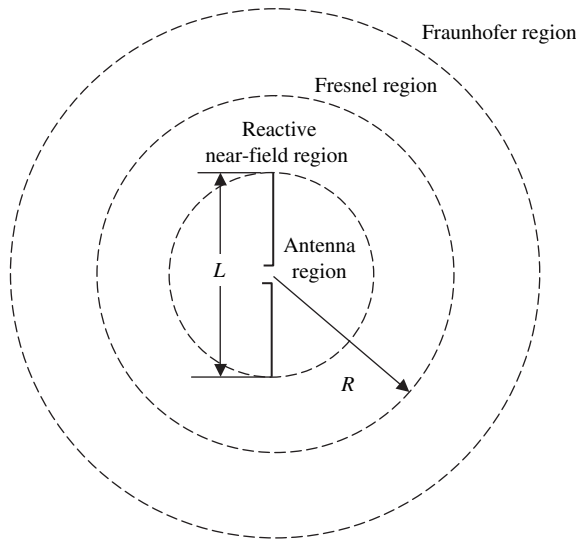


Figure 3.1 Antenna field regions.

The boundaries defined in the regions above are not arbitrarily assigned but are the consequence of solving the exact fields surrounding a finite length antenna.

The four regions and their boundaries are defined as follows:

Antenna region That region which circumscribes the physical antenna boundaries is called the *antenna region* as defined by

$$R \leq \frac{L}{2}$$

Reactive near-field region That region which contains the reactive energy surrounding the antenna is called the *reactive near-field region*. It represents energy stored in the vicinity of the antenna which does not radiate and thus is seen in the imaginary part of the antenna terminal impedance. This region is defined by

$$R \leq 0.62 \sqrt{\frac{L^3}{\lambda}}$$

Fresnel region (radiating near field) That region which lies between the reactive near-field and the Fraunhofer far-field is the *Fresnel region* or *radiating near-field region*. The antenna field radiates in this region but the radiation pattern changes with distance from the phase center owing to the fact that the radiated field components

diminish at different rates. This region is defined by

$$0.62\sqrt{\frac{L^3}{\lambda}} \leq R \leq \frac{2L^2}{\lambda}$$

Fraunhofer region (far-field) That region which lies beyond the near-field and where the radiation pattern is unchanging with distance is defined as the *Fraunhofer region*. This is the principle region of operation for most elemental antennas. This region is defined by

$$R \geq \frac{2L^2}{\lambda}$$

For practical purposes, this text generally assumes antenna radiation in the Fresnel or Fraunhofer regions. If array element coupling is to be considered, the reactive near-field region must be accounted for in all calculations.

3.2 Power Density

All radiated antenna fields carry power away from the antenna which can be intercepted by distant receiving antennas. It is this power which is of use in communication systems. As a trivial example, let us assume that the propagating phasor fields, generated by a point source isotropic antenna, are given further and are expressed in spherical coordinates.

$$\bar{E}_{\theta s} = \frac{E_0}{r} e^{-jkr} \hat{\theta} \text{ V/m} \quad (3.1)$$

$$\bar{H}_{\phi s} = \frac{E_0}{\eta r} e^{-jkr} \hat{\phi} \text{ A/m} \quad (3.2)$$

where η is the intrinsic impedance of the medium.

If the intrinsic medium is lossless, the time-varying instantaneous fields can be easily derived from Eqs. (3.1) and (3.2) to be

$$\bar{E}(r, t) = \text{Re} \left\{ \frac{E_0}{r} e^{j(\omega t - kr)} \hat{\theta} \right\} = \frac{E_0}{r} \cos(\omega t - kr) \hat{\theta} \quad (3.3)$$

$$\bar{H}(r, t) = \text{Re} \left\{ \frac{E_0}{\eta r} e^{j(\omega t - kr)} \hat{\phi} \right\} = \frac{E_0}{\eta r} \cos(\omega t - kr) \hat{\phi} \quad (3.4)$$

The electric field intensity in Eq. (3.3) is seen to radiate in the positive r direction and is polarized in the positive $\hat{\theta}$ direction. The magnetic field intensity in Eq. (3.4) is seen to radiate in the positive r direction and is polarized in the positive $\hat{\phi}$ direction. Figure 3.2 shows the field vectors

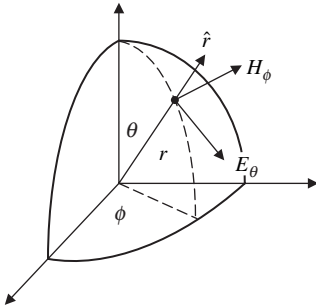


Figure 3.2 EM field radiating from a point source.

in spherical coordinates. These far-fields are mutually perpendicular and are tangent to the sphere whose radius is r .

The Poynting vector, named after J. H. Poynting,¹ is the cross product of the electric and magnetic field intensities and is given as

$$\bar{P} = \bar{E} \times \bar{H} \text{ W/m}^2 \quad (3.5)$$

The cross product is in the right-handed sense and gives the direction of propagation of the power density. The Poynting vector is a measure of the instantaneous power density flow away from the source. By substituting Eqs. (3.3) and (3.4) into Eq. (3.5) and using a simple trigonometric identity we get

$$\bar{P}(r, t) = \frac{E_0^2}{2\eta r^2} [1 + \cos(2\omega t - 2kr)] \hat{r} \quad (3.6)$$

The first term in Eq. (3.6) represents the time average power density radiating away from the antenna whereas the second term represents an instantaneous ebb and flow. By taking the time average of Eq. (3.6) we can define the average power density.

$$\begin{aligned} \bar{W}(r) &= \frac{1}{T} \int_0^T \bar{P}(r, t) dt \\ &= \frac{E_0^2}{2\eta r^2} \hat{r} \text{ W/m}^2 \end{aligned} \quad (3.7)$$

The calculation of the time average power density is equivalent to performing a calculation in phasor space.

$$\bar{W}(r, \theta, \phi) = \frac{1}{2} \text{Re}(\bar{E}_s \times \bar{H}_s^*) = \frac{1}{2\eta} |\bar{E}_s|^2 \quad (3.8)$$

¹John Henry Poynting (1852–1914). A student of Maxwell who derived an equation to show energy flow from EM fields.

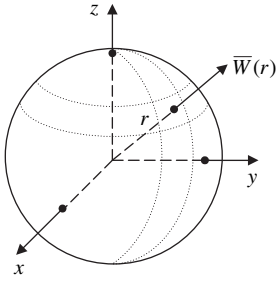


Figure 3.3 Power density from an isotropic point source.

Equation (3.8) represents the average power density flow away from the isotropic antenna and thus is not a function of θ or ϕ . For practical antennas, the power density is always a function of r and at least one angular coordinate.

In general, the power density can be represented as a power flow through a sphere of radius r as shown in Fig. 3.3.

The total power radiated by an antenna is found by the closed surface integral of the power density over the sphere bounding the antenna. This is equivalent to applying the divergence theorem to the power density. The total power is thus given by

$$\begin{aligned} P_{tot} &= \oint \bar{W} \cdot d\bar{s} = \int_0^{2\pi} \int_0^\pi W_r(r, \theta, \phi) r^2 \sin \theta d\theta d\phi \\ &= \int_0^{2\pi} \int_0^\pi W_r(r, \theta, \phi) r^2 d\Omega \quad W \end{aligned} \quad (3.9)$$

where $d\Omega = \sin \theta d\theta d\phi$ = element of solid angle or the differential solid angle.

In the isotropic case where the power density is not a function of θ or ϕ , Eq. (3.9) simplifies to become

$$P_{tot} = \int_0^{2\pi} \int_0^\pi W_r(r) r^2 \sin \theta d\theta d\phi = 4\pi r^2 W_r(r) \quad (3.10)$$

or conversely

$$W_r(r) = \frac{P_{tot}}{4\pi r^2} \quad (3.11)$$

Thus, for isotropic antennas, the power density is found by uniformly spreading the total power radiated over the surface of a sphere of radius r . Thus the density diminishes inversely with r^2 . It is also interesting to observe that the power density is only a function of the real power (P_{tot}) delivered to the antenna terminals. The reactive power does not contribute to the radiated fields.

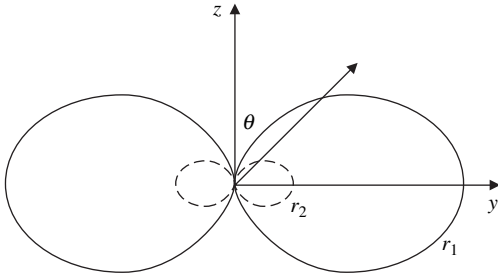


Figure 3.4 Pattern plots for two ranges r_1 and r_2 .

Example 3.1 Find the total power radiated by an isotropic antenna whose electric field intensity is given as

$$\vec{E}_s = \frac{2}{r} e^{-jkr} \hat{\theta} \text{ V/m}$$

Solution Equation (3.7) shows the power density to be

$$\bar{W}(r) = \frac{5.3 \times 10^{-3}}{r^2} \hat{r} \text{ W/m}^2$$

Substituting this result into Eq. (3.9) yields the total power to be

$$P_{tot} = 66.7 \text{ mW}$$

Example 3.2 Plot the power density vs. angle for the far electric field at two different distances r_1 and r_2 . Where $r_1 = 100 \text{ m}$ and $r_2 = 200 \text{ m}$. Given that

$$\vec{E}_s = \frac{100 \sin \theta}{r} e^{-jkr} \hat{\theta}$$

$$\vec{H}_s = \frac{100 \sin \theta}{\eta r} e^{-jkr} \hat{\phi}$$

Solution By using Eq. (3.8), the power density magnitude can easily be found to be

$$W(r, \theta) = \frac{13.3 \sin^2 \theta}{r^2}$$

Using MATLAB code `sa_ex3_2.m` and the polar plot command we get a pattern plot for both distances as shown in Fig. 3.4.

3.3 Radiation Intensity

The radiation intensity can be viewed as a distance normalized power density. The power density in Eq. (3.8) is inversely proportional to the distance squared and thus diminishes rapidly moving away from the antenna. This is useful in indicating power levels but is not useful in

indicating distant antenna patterns. The radiation intensity removes the $1/r^2$ dependence thus making far-field pattern plots distance independent. The radiation intensity is thus defined as

$$U(\theta, \phi) = r^2 |\overline{W}(r, \theta, \phi)| = r^2 W_r(r, \theta, \phi) \tag{3.12}$$

It is easy to show Eq. (3.12) can alternatively be expressed by

$$\begin{aligned} U(\theta, \phi) &= \frac{r^2}{2\eta} |\overline{E}_s(r, \theta, \phi)|^2 \\ &= \frac{\eta r^2}{2} |\overline{H}_s(r, \theta, \phi)|^2 \end{aligned} \tag{3.13}$$

This definition also simplifies the calculation of the total power radiated by the antenna. Equation (3.9) can be repeated substituting the radiation intensity.

$$\begin{aligned} P_{tot} &= \int_0^{2\pi} \int_0^\pi W_r(r, \theta, \phi) r^2 \sin \theta d\theta d\phi \\ &= \int_0^{2\pi} \int_0^\pi U(\theta, \phi) d\Omega \quad W \end{aligned} \tag{3.14}$$

The general radiation intensity indicates the radiation pattern of the antenna in three dimensions. All anisotropic antennas have a nonuniform radiation intensity and therefore a nonuniform radiation pattern. Figure 3.5 shows an example of a three-dimensional pattern as displayed in spherical coordinates. This antenna pattern or *beam* pattern is an indication of the directions in which the signal is radiated. In the

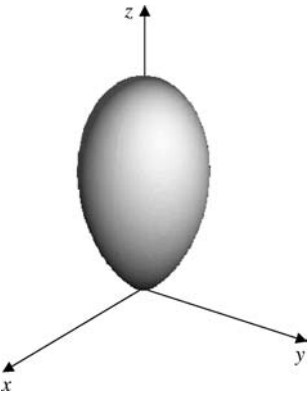


Figure 3.5 Antenna 3-D pattern.

case of Fig. 3.5, the maximum radiation is in the $\theta = 0$ direction or along the z axis.

Example 3.3 In the Fraunhofer region (far-field), a small dipole has an electric field intensity given by

$$\vec{E}_s(r, \theta, \phi) = \frac{E_0 \sin \theta}{r} e^{-jkr} \hat{\theta}$$

What are the radiation intensity and the total power radiated by this antenna?

Solution Using Eq. (3.13), the radiation intensity is given as,

$$U(\theta, \phi) = \frac{E_0^2 \sin^2 \theta}{2\eta}$$

Using Eq. (3.14), the total power radiated is given as,

$$P_{tot} = 0.011 E_0^2 \text{ W}$$

3.4 Basic Antenna Nomenclature

With an understanding of the derivation of the radiation intensity ($U(\theta, \phi)$), we can now define some metrics that help define an antenna's performance.

3.4.1 Antenna pattern

An antenna pattern is either a function or a plot describing the directional properties of an antenna. The pattern can be based upon the function describing the electric or magnetic fields. In that case, the pattern is called a *field pattern*. The pattern can also be based upon the radiation intensity function defined in the previous section. In that case, the pattern is called a *power pattern*. The antenna pattern may not come from a functional description but also may be the result of antenna measurements. In this case the measured pattern can be expressed as a *field pattern* or as a *power pattern*.

Figures 3.6a and b show a typical two-dimensional field pattern plot displayed in both rectangular coordinates and in polar coordinates. The mainlobe and sidelobes of the pattern are indicated. The mainlobe is that portion of the pattern which has maximum intended radiation. The sidelobes are generally unintended radiation directions.

Figures 3.6a and b can be viewed as demonstrating a two-dimensional slice of what is typically a three-dimensional pattern. Figure 3.7 shows the same pattern displayed in three dimensions. The

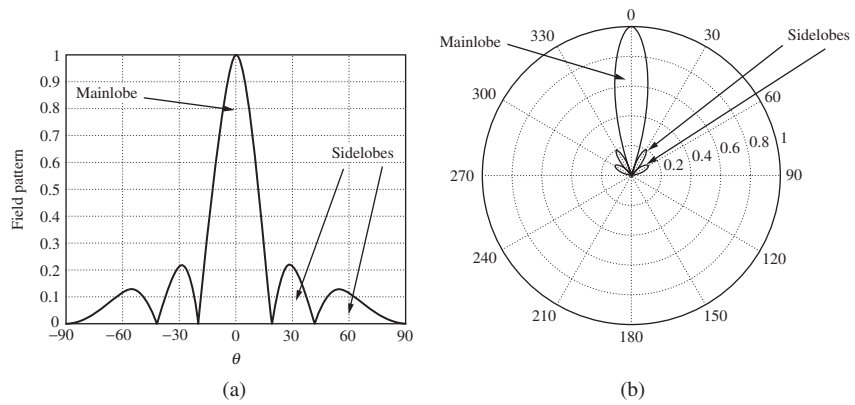


Figure 3.6 (a) Field pattern plot in rectangular coordinates. (b) Field pattern plot in polar coordinates.

three-dimensional perspective is useful for illustration purposes but often antenna designers display two-dimensional plots in the principle planes of three-dimensional patterns.

Example 3.4 Use MATLAB to produce a three-dimensional radiation pattern for a radiation intensity given by $U(\theta) = \cos^2 \theta$.

Solution Radiation patterns are normally calculated in spherical coordinates. That being understood, the patterns have corresponding x , y , and z coordinate values. We can simply use the coordinate transformation to transform points on the radiation intensity to (x, y, z) coordinates. Thus

$$x = U(\theta, \phi) \sin \theta \cos \phi \qquad y = U(\theta, \phi) \sin \theta \sin \phi \qquad z = U(\theta, \phi) \cos \theta$$

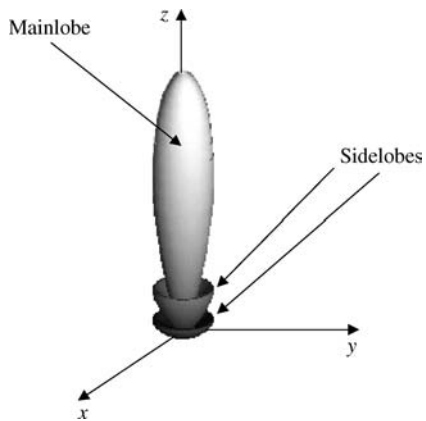


Figure 3.7 Field pattern plot in three dimensions.

MATLAB has a feature called *ezmesh* which allows us to easily plot a surface in three dimensions. The corresponding MATLAB commands are

```
fx = inline('cos(theta)^2*sin(theta)*cos(phi)')
fy = inline('cos(theta)^2*sin(theta)*sin(phi)')
fz = inline('cos(theta)^2*cos(theta)')
figure
ezmesh(fx,fy,fz,[0 2*pi 0 pi],100)
colormap([0 0 0])
axis equal
set(gca,'xdir','reverse','ydir','reverse')
```

The plot from MATLAB code `sa_ex3_4.m` is shown in Fig. 3.8.

3.4.2 Antenna boresight

The antenna boresight is the intended physical aiming direction of an antenna or the direction of maximum gain. This is also the central axis of the antenna's mainlobe. In other words, it is the normally intended direction for maximum radiation. The boresight in Fig. 3.7 is central axis of the mainlobe which corresponds to the z -axis where $\theta = 0$.

3.4.3 Principal plane patterns

The field patterns or the power patterns are usually taken as two-dimensional slices of the three-dimensional antenna pattern. These

$$x = \cos(\theta)^2 \sin(\theta) \cos(\phi) \quad y = \cos(\theta)^2 \sin(\theta) \sin(\phi) \quad z = \cos(\theta)^2 \cos(\theta)$$

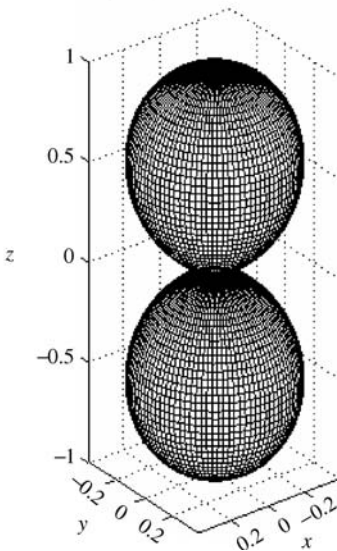


Figure 3.8 3-D radiation pattern.

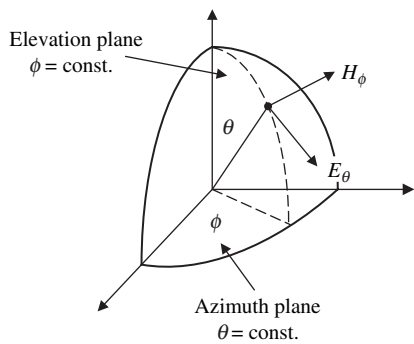


Figure 3.9 Principle planes in spherical coordinates.

slices can be defined in several different ways. One option is to plot the E- and H-plane patterns when the radiation is linearly polarized. In this case, the E-plane pattern is the plot containing the E field vector and the direction of maximum radiation. In the other case, the H-plane pattern is the plot containing the H field vector and the direction of maximum radiation. It is most convenient to orient the antenna in spherical coordinates such that the E- and H-planes correspond to the θ and ϕ constant planes. These planes are called the *azimuth* and *elevation* planes, respectively. Figure 3.9 shows the spherical coordinate system with polarized fields and the azimuth and elevation planes parallel to these field vectors.

3.4.4 Beamwidth

The beamwidth is measured from the 3-dB points of a radiation pattern. Figure 3.10 shows a two-dimensional slice of Fig. 3.5. The beamwidth is the angle between the 3-dB points. Since this is a power pattern, the 3-dB points are also the half power points.

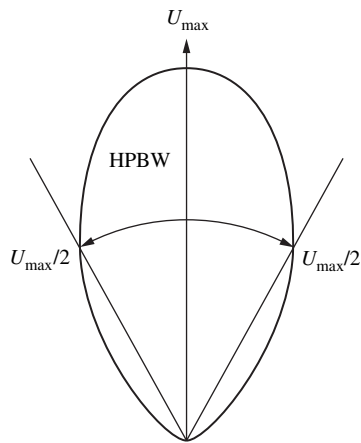


Figure 3.10 Half-power beamwidth.

In the case of field patterns instead of power patterns, the 3-dB points would be when the normalized pattern amplitude $= \frac{1}{\sqrt{2}} = 0.707$.

3.4.5 Directivity

The directivity is a measure of how directive an individual antenna is relative to an isotropic antenna radiating the same total power. In other words, the directivity is the ratio of the power density of an anisotropic antenna relative to an isotropic antenna radiating the same total power. Thus the directivity is given as

$$D(\theta, \phi) = \frac{W(\theta, \phi)}{\frac{P_{tot}}{4\pi r^2}} = \frac{4\pi U(\theta, \phi)}{P_{tot}} \quad (3.15)$$

The directivity can be more explicitly expressed by substituting Eq. (3.14) into (3.15) to get

$$D(\theta, \phi) = \frac{4\pi U(\theta, \phi)}{\int_0^{2\pi} \int_0^\pi U(\theta, \phi) \sin \theta d\theta d\phi} \quad (3.16)$$

The maximum directivity is a constant and is simply the maximum of Eq. (3.16). The maximum directivity is normally denoted by D_0 . Thus, the maximum directivity is found by a slight modification of Eq. (3.16) to be

$$D_0 = \frac{4\pi U_{max}}{\int_0^{2\pi} \int_0^\pi U(\theta, \phi) \sin \theta d\theta d\phi} \quad (3.17)$$

The directivity of an isotropic source is always equal to 1 since isotropic sources radiate equally in all directions and therefore are not inherently directive.

Example 3.5 Find the directivity and the maximum directivity of the small-dipole radiation pattern given in Example 3.3.

Solution Using Eq. (3.16) we find

$$D(\theta, \phi) = \frac{4\pi \sin^2 \theta}{\int_0^{2\pi} \int_0^\pi \sin^3 \theta d\theta d\phi} = 1.5 \sin^2 \theta$$

As can intuitively be seen in the Example 3.5, the directivity is not affected by the radiation intensity amplitude but is only dependent upon the functional form of $U(\theta, \phi)$. The scalar amplitude terms divide out. The maximum directivity is a constant and is simply given by Eq. (3.17) to be $D_0 = 1.5$.

Plotting the directivity of an antenna is more useful than plotting the radiation intensity because the amplitude indicates the performance relative to an isotropic radiator irrespective of the distance. This is helpful in indicating not just the antenna pattern but also indicating a form of the antenna gain.

3.4.6 Beam solid angle

The beam solid angle (Ω_A) is that angle through which all of the antenna power radiates if its radiation intensity were equal to its maximum radiation intensity ($U_{\max}$). The beam solid angle can be seen by expressing Eq. (3.17) as

$$D_0 = \frac{4\pi}{\int_0^{2\pi} \int_0^\pi \frac{U(\theta, \phi)}{U_{\max}} \sin \theta d\theta d\phi} = \frac{4\pi}{\Omega_A} \quad (3.18)$$

where

$$\Omega_A = \text{Beam solid angle} = \int_0^{2\pi} \int_0^\pi \frac{U(\theta, \phi)}{U_{\max}} \sin \theta d\theta d\phi \quad (3.19)$$

The beam solid angle is given in *steradians* where one steradian is defined as the solid angle of a sphere subtending an area on the surface of the sphere equal to r^2 . Thus there are 4π steradians in a sphere. The beam solid angle is the spatial version of the equivalent noise bandwidth in communications. An explanation of the noise equivalent bandwidth is given by Haykin [4].

3.4.7 Gain

The directivity of an antenna is an indication of the directionality of an antenna. It is the ability of an antenna to direct energy in preferred directions. The directivity assumes that there are no antenna losses through conduction losses, dielectric losses, and transmission line mismatches. The antenna gain is a modification of the directivity so as to include the effects of antenna inefficiencies. The gain is more reflective of an actual antenna's performance. The antenna gain expression is given by

$$G(\theta, \phi) = eD(\theta, \phi) \quad (3.20)$$

where e is the total antenna efficiency including the effects of losses and mismatches. The pattern produced by the gain is identical to the pattern produced by the directivity except for the efficiency scale factor e .

3.4.8 Effective aperture

Just as an antenna can radiate power in various preferred directions it can also receive power from the same preferred directions. This principle is called *reciprocity*. Figure 3.11 shows transmit and receive antennas. The transmit antenna is transmitting with power P_1 (Watts) and radiates a power density W_1 (Watts/m²).

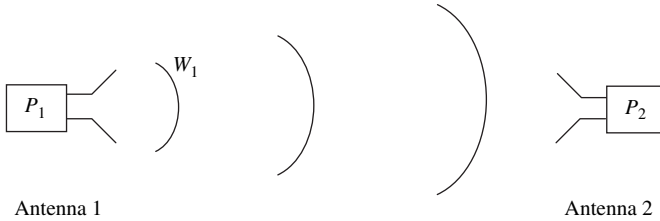


Figure 3.11 Transmit and receive antennas.

The receive antenna intercepts a portion of the incident power density W_1 thereby delivering power P_2 to the load. The receive antenna can be viewed as an effective aperture of area A_{e2} which captures a portion of the available power density. Thus, using Eqs. (3.15) and (3.20), we can write the received power as

$$P_2 = A_{e2} W_1 = \frac{A_{e2} P_1 e_1 D_1(\theta_1, \phi_1)}{4\pi r_1^2} W \quad (3.21)$$

where r_1, θ_1, ϕ_1 are the local spherical coordinates for antenna 1.

If the antennas in Fig. 3.10 are reversed such that the receive antenna transmits and the transmit antenna receives, it can be shown that

$$P_1 = A_{e1} W_2 = \frac{A_{e1} P_2 e_2 D_2(\theta_2, \phi_2)}{4\pi r_2^2} W \quad (3.22)$$

where r_2, θ_2, ϕ_2 are the local spherical coordinates for antenna 2.

The derivation is beyond the scope of this text but it can be shown from Eqs. (3.21) and (3.22) [1,2] that the effective aperture is related to the directivity of an antenna by

$$A_e(\theta, \phi) = \frac{\lambda^2}{4\pi} e D(\theta, \phi) = \frac{\lambda^2}{4\pi} G(\theta, \phi) \quad (3.23)$$

3.5 Friis Transmission Formula

Harald Friis² devised a formula relating the transmit and received powers between two distant antennas. We will assume that the transmit and receive antennas are polarization matched. (That is, the polarization of the receive antenna is perfectly matched to the polarization created by the transmit antenna.) By substituting Eq. (3.23) into (3.21)

²Harald T. Friis (1893–1976): In 1946 developed a transmission formula.

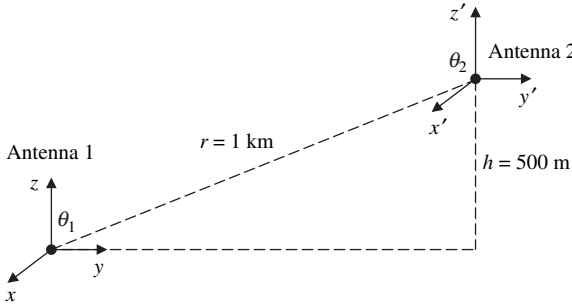


Figure 3.12 Two antennas separated by a distance r .

one can derive the following relationship:

$$\frac{P_2}{P_1} = \left(\frac{\lambda}{4\pi r} \right)^2 G_1(\theta_1, \phi_1) G_2(\theta_2, \phi_2) \quad (3.24)$$

If polarization must be taken into account, we can multiply Eq. (3.24) by the polarization loss factor (PLF) = $|\hat{\rho}_1 \cdot \hat{\rho}_2|^2$ where $\hat{\rho}_1$ and $\hat{\rho}_2$ are the polarizations of antennas 1 and 2, respectively. An intensive treatment of polarization and the polarization loss factor can be found in ref. [2].

Example 3.6 Calculate the receive power P_2 at antenna 2 if the transmit power $P_1 = 1$ kW. The transmitter gain is $G_1(\theta_1, \phi_1) = \sin^2(\theta_1)$ and the receiver gain is $G_2(\theta_2, \phi_2) = \sin^2(\theta_2)$. The operating frequency is 2 GHz. Use the Fig. 3.12 to solve the problem. The y - z and y' - z' planes are coplanar.

Solution The geometry indicates $\theta_1 = 60^\circ$ and $\theta_2 = 120^\circ$. Using Eq. (3.24) we have

$$\begin{aligned} P_2 &= P_1 \left(\frac{\lambda}{4\pi r} \right)^2 G_1(\theta_1, \phi_1) G_2(\theta_2, \phi_2) \\ &= P_1 \left(\frac{\lambda}{4\pi r} \right)^2 \sin^2(\theta_1) \sin^2(\theta_2) \\ &= 10^3 \left(\frac{15 \text{ cm}}{4\pi 10^3} \right)^2 \sin^2(60) \sin^2(120) \\ &= 80.1 \text{ nW} \end{aligned}$$

3.6 Magnetic Vector Potential and the Far Field

All antennas radiate electric and magnetic fields by virtue of charges accelerating on the antenna. These accelerating charges are normally in the form of an ac current. While it is possible to calculate distant

radiated fields directly from the antenna currents, in most cases it is mathematically unwieldy. Therefore an intermediate step is used where we instead calculate the magnetic vector potential. From the vector potential we can then find the distant radiated fields.

The point form of Gauss' law states that $\nabla \cdot \bar{B} = 0$. Since the divergence of the curl of any vector is always identically equal to zero ($\nabla \cdot \nabla \times \bar{A} = 0$), we then can define the $\bar{B}$ field in terms of the magnetic vector potential $\bar{A}$.

$$\bar{B} = \nabla \times \bar{A} \quad (3.25)$$

Equation (3.25) solves Gauss' law.

Since $\bar{B}$ and $\bar{H}$ are related by a constant, we can also write

$$\bar{H} = \frac{1}{\mu} \nabla \times \bar{A} \quad (3.26)$$

The electric field, in a source free region, can be derived from the magnetic field by

$$\bar{E} = \frac{1}{j\omega\epsilon} \nabla \times \bar{H} \quad (3.27)$$

Thus, if we know the vector potential, we can subsequently calculate $\bar{E}$ and $\bar{H}$ fields.

Figure 3.13 shows an arbitrary current source $\bar{I}$ creating a distant vector potential $\bar{A}$.

The vector potential is related to the current source by

$$\bar{A} = \frac{\mu}{4\pi} \int \bar{I}(r') \frac{e^{-jkR}}{R} dl' \quad (3.28)$$

where $\bar{I}(r') = I_x(r')\hat{x} + I_y(r')\hat{y} + I_z(r')\hat{z}$ = current in three dimensions

$\bar{r}'$ = position vector in source coordinates

$\bar{r}$ = position vector in field coordinates

$\bar{R}$ = distance vector = $\bar{r} - \bar{r}'$

$R = |\bar{R}|$

dl' = differential length at current source

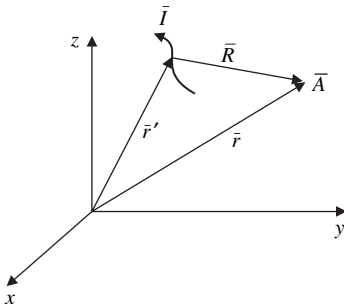


Figure 3.13 Current source and distant vector potential.

The vector potential can be found from any line current source using Eq. (3.28). The results can be substituted into Eqs. (3.26) and (3.27) to find the distant fields. Two of the easier antenna problems to solve are the linear antenna and the loop antenna.

3.7 Linear Antennas

Fundamental to understanding antenna radiation is to understand the behavior of a straight wire or linear antenna. Not only is the math greatly simplified by a straight wire segment, but the linear antenna solution gives insight into the behavior of many more complicated structures which can often be viewed as a collection of straight wire segments.

3.7.1 Infinitesimal dipole

The infinitesimal dipole is a short wire segment antenna where the length $L \ll \lambda$. It is aligned along the z -axis symmetrically placed about the x - y plane as shown in Fig. 3.14.

The phasor current is given by $\bar{I} = I_0 \hat{z}$. The position and distance vectors are given by $\bar{r} = r\hat{r} = x\hat{x} + y\hat{y} + z\hat{z}$, $\bar{r}' = z'\hat{z}$, and $\bar{R} = x\hat{x} + y\hat{y} + (z - z')\hat{z}$. Thus, the vector potential is given by

$$\bar{A} = \frac{\mu_0}{4\pi} \int_{-L/2}^{L/2} I_0 \hat{z} \frac{e^{-jk\sqrt{x^2+y^2+(z-z')^2}}}{\sqrt{x^2+y^2+(z-z')^2}} dz' \quad (3.29)$$

Since we are assuming an infinitesimal dipole ($r \gg z'$), then $R \approx r$. Thus, the integral can easily be solved yielding

$$\bar{A} = \frac{\mu_0}{4\pi} \int_{-L/2}^{L/2} I_0 \hat{z} \frac{e^{-jkr}}{r} dz' = \frac{\mu_0 I_0 L}{4\pi r} e^{-jkr} \hat{z} = A_z \hat{z} \quad (3.30)$$

Since most antenna fields are more conveniently expressed in spherical coordinates, we may apply a vector transformation to Eq. (3.30) or we may graphically determine $\bar{A}$ in spherical coordinates. Figure 3.15 shows the vector potential in both rectangular and spherical coordinates.

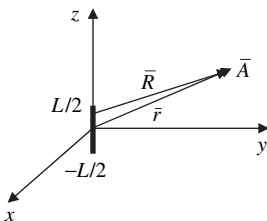


Figure 3.14 Infinitesimal dipole.

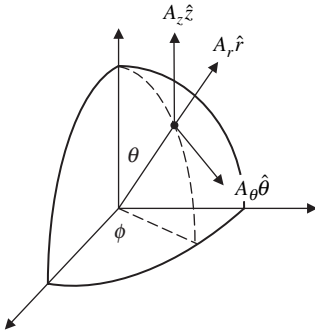


Figure 3.15 Vector potential for an infinitesimal dipole.

The A_r and A_θ components are vector projections of A_z onto the corresponding axes. It can thus be shown that

$$A_r = A_z \cos \theta = \frac{\mu_0 I_0 L e^{-jkr}}{4\pi r} \cos \theta \quad (3.31)$$

$$A_\theta = -A_z \sin \theta = -\frac{\mu_0 I_0 L e^{-jkr}}{4\pi r} \sin \theta \quad (3.32)$$

Since $\vec{H}$ is related to the curl of $\vec{A}$ in Eq. (3.26), we can take the curl in spherical coordinates to get

$$H_\phi = \frac{jkI_0 L \sin \theta}{4\pi r} \left[1 + \frac{1}{jkr} \right] e^{-jkr} \quad (3.33)$$

where $H_r = 0$ and $H_\theta = 0$.

The electric field can be found from Eq. (3.33) by substituting into Eq. (3.27) giving

$$E_r = \frac{\eta I_0 L \cos \theta}{2\pi r^2} \left[1 + \frac{1}{jkr} \right] e^{-jkr} \quad (3.34)$$

$$E_\theta = \frac{jk\eta I_0 L \sin \theta}{4\pi r} \left[1 + \frac{1}{jkr} - \frac{1}{(kr)^2} \right] e^{-jkr} \quad (3.35)$$

where $E_\phi = 0$ and η = intrinsic impedance of the medium.

In the far-field, the higher order terms involving $1/r^2$ and $1/r^3$ become negligibly small simplifying Eqs. (3.34) and (3.35) to become

$$E_\theta = \frac{jk\eta I_0 L \sin \theta}{4\pi r} e^{-jkr} \quad (3.36)$$

$$H_\phi = \frac{jkI_0 L \sin \theta}{4\pi r} e^{-jkr} \quad (3.37)$$

It should be noted that in the far field $\frac{E_\theta}{H_\phi} = \eta$.

Power density and radiation intensity. We can calculate the far field power density and the radiation intensity by substituting Eq. (3.36) into Eq. (3.8) and (3.12) to get

$$W_r(\theta, \phi) = \frac{1}{2\eta} \left| \frac{k\eta I_0 L \sin \theta}{4\pi r} \right|^2 = \frac{\eta}{8} \left| \frac{I_0 L}{\lambda} \right|^2 \frac{\sin^2 \theta}{r^2} \quad (3.38)$$

$$U(\theta) = \frac{\eta}{8} \left| \frac{I_0 L}{\lambda} \right|^2 \sin^2 \theta \quad (3.39)$$

where I_0 = complex phasor current = $|I_0|e^{j\psi}$
 λ = wavelength.

Figure 3.16 shows a plot of the normalized radiation intensity given in Eq. (3.39). The radiation intensity is superimposed on a rectangular coordinate system. Since the antenna is aligned along the z -axis, the maximum radiation is broadside to the infinitesimal dipole.

Directivity. The directivity, as defined in Eq. (3.16), can be applied to the infinitesimal dipole radiation intensity. The constant terms divide out in the numerator and denominator yielding

$$D(\theta) = \frac{4\pi \sin^2 \theta}{\int_0^{2\pi} \int_0^\pi \sin^3 \theta d\theta d\phi} = 1.5 \sin^2 \theta \quad (3.40)$$

This solution is exactly the same as derived in Example 3.3.

3.7.2 Finite length dipole

The same procedure can also be applied to a finite length dipole to determine the far fields and the radiation pattern. However, since a finite length dipole can be viewed as the concatenation of numerous infinitesimal dipoles, we can use the principle of superposition to find

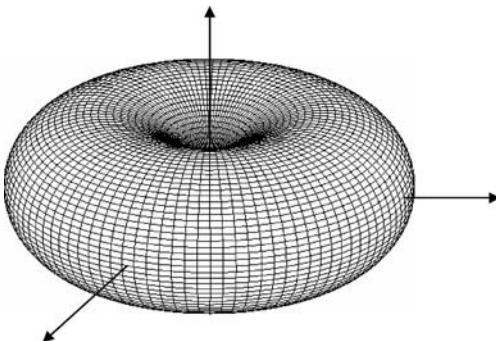


Figure 3.16 3-D plot of infinitesimal dipole radiation.

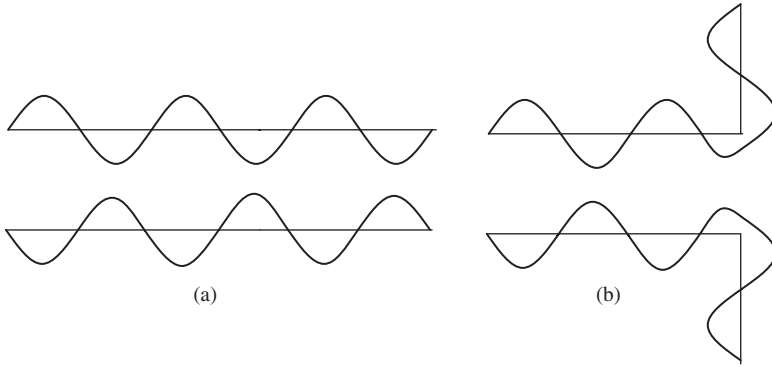


Figure 3.17 Standing waves on a transmission line and on a dipole.

the fields. Superposing numerous infinitesimal dipoles of length dz' results in an integral given as [1]

$$E_{\theta} = \frac{j\eta k e^{-jkr}}{4\pi r} \sin \theta \int_{-L/2}^{L/2} I(z') e^{jkz' \cos \theta} dz' \quad (3.41)$$

Since dipoles are center fed and the currents must terminate at the ends, a good approximation for the dipole current is given as sinusoidal (King [5]). It is well known that a twin lead transmission line with an open circuit termination generates sinusoidal standing waves along the conductors. If the ends of the leads are bent, so as to form a dipole, the currents can still be approximated as piece-wise sinusoidal. Figure 3.17a shows a twin lead transmission line with sinusoidal currents. Figure 3.17b shows a twin lead transmission line with sinusoidal currents which is terminated into a dipole. The dipole currents can be viewed as an extension of the existing transmission line currents.

Since sinusoidal currents are a good approximation to the currents on a linear antenna, we can devise an analytic expression for the current in Eq. (3.41) ([1,3]) as

$$I(z') = \begin{cases} I_0 \sin[k(\frac{L}{2} - z')] & 0 \leq z' \leq L/2 \\ I_0 \sin[k(\frac{L}{2} + z')] & -L/2 \leq z' \leq 0 \end{cases} \quad (3.42)$$

By substituting Eq. (3.42) into Eq. (3.41) we can solve for the approximate electric far field as

$$E_{\theta} = \frac{j\eta I_0 e^{-jkr}}{2\pi r} \left[\frac{\cos(\frac{kL}{2} \cos \theta) - \cos(\frac{kL}{2})}{\sin \theta} \right] \quad (3.43)$$

The magnetic field can easily be found to be

$$H_\phi = \frac{jI_0 e^{-jkr}}{2\pi r} \left[\frac{\cos\left(\frac{kL}{2} \cos \theta\right) - \cos\left(\frac{kL}{2}\right)}{\sin \theta} \right] \quad (3.44)$$

Power density and radiation intensity. We can again calculate the far field power density and the radiation intensity to be

$$W_r(\theta, \phi) = \frac{1}{2\eta} |E_\theta|^2 = \frac{\eta}{8} \left| \frac{I_0}{\pi r} \right|^2 \left[\frac{\cos\left(\frac{kL}{2} \cos \theta\right) - \cos\left(\frac{kL}{2}\right)}{\sin \theta} \right]^2 \quad (3.45)$$

$$U(\theta) = \frac{\eta}{8} \left| \frac{I_0}{\pi} \right|^2 \left[\frac{\cos\left(\frac{kL}{2} \cos \theta\right) - \cos\left(\frac{kL}{2}\right)}{\sin \theta} \right]^2 \quad (3.46)$$

Figure 3.18 shows plots of the normalized radiation intensity given in Eq. (3.46) in three dimensions for $\frac{L}{\lambda} = 0.5, 1$, and 1.5 . It can be seen that as the dipole length increases, the mainlobe narrows. However, in the $L/\lambda = 1.5$ case, the mainlobe no longer is perpendicular to the dipole axis. Usually dipoles are designed to be of length $L = \lambda/2$.

Directivity. The directivity of the finite length dipole is given by substituting Eq. (3.46) into Eq. (3.16).

$$D(\theta) = \frac{4\pi \left[\frac{\cos\left(\pi \frac{L}{\lambda} \cos \theta\right) - \cos\left(\pi \frac{L}{\lambda}\right)}{\sin \theta} \right]^2}{\int_0^{2\pi} \int_0^\pi \left[\frac{\cos\left(\pi \frac{L}{\lambda} \cos \theta\right) - \cos\left(\pi \frac{L}{\lambda}\right)}{\sin \theta} \right]^2 \sin \theta d\theta d\phi} \quad (3.47)$$

where L/λ is the length in wavelengths.

The maximum directivity is given when the numerator is at its maximum value. We can plot the finite dipole maximum directivity (D_0) vs. the length L in wavelengths to yield Fig. 3.19.

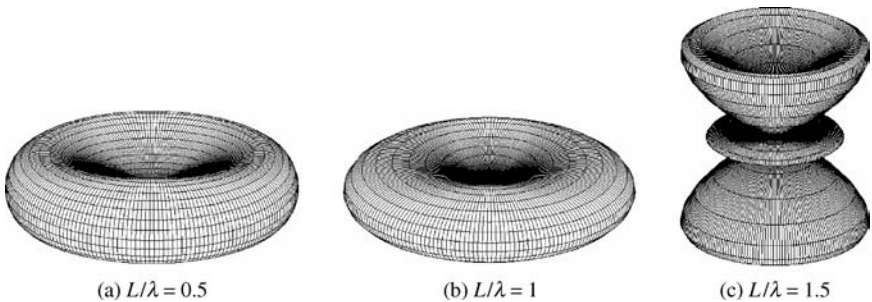


Figure 3.18 Finite dipole radiation intensity.

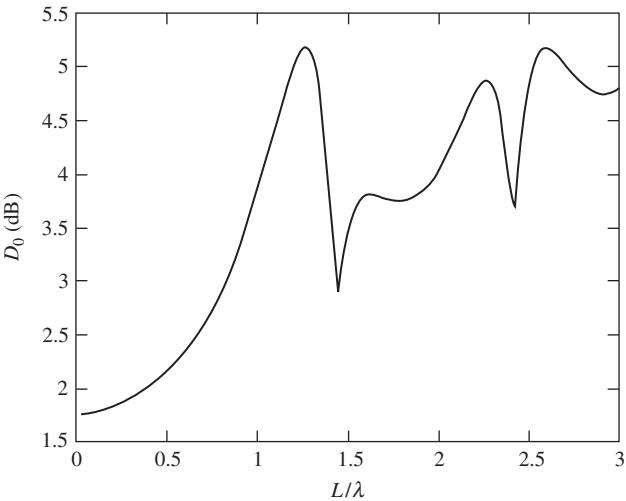


Figure 3.19 Maximum directivity of finite length dipole.

3.8 Loop Antennas

In addition to solving for the fields radiated from linear antennas, it is also instructive to calculate fields radiated by loop antennas. Both linear and loop antennas form the backbone for numerous antenna problems and both forms of antennas shed light on general antenna behavior.

3.8.1 Loop of constant phasor current

Figure 3.20 shows a loop antenna of radius a , centered on the z axis and residing in the x - y plane. The loop current I_0 flows in the $\hat{\phi}'$ direction.

We will assume the far field condition such that $(a \ll r)$. We can therefore assume that $\bar{r}$ and $\bar{R}$ are approximately parallel to each other.

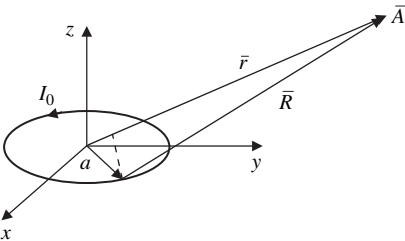


Figure 3.20 Small loop antenna.

We can now more easily determine the vector $\bar{R}$. We can define the following variables:

$$\bar{r} = r \sin \theta \cos \phi \hat{x} + \sin \theta \sin \phi \hat{y} + \cos \theta \hat{z} \quad (3.48)$$

$$\bar{r}' = a \cos \phi' \hat{x} + a \sin \phi' \hat{y} \quad (3.49)$$

$$R \approx r - \bar{r}' \cdot \hat{r} = r - a \sin \theta \cos(\phi - \phi') \quad (3.50)$$

The vector potential is given by the integral of Eq. (3.27) and is found to be

$$\bar{A} = \frac{\mu_0}{4\pi} \int_0^{2\pi} I_0 \hat{\phi}' \frac{e^{-jkR}}{R} a d\phi' \quad (3.51)$$

where $\hat{\phi}'$ is the current direction in source coordinates. It can be converted to field coordinates by the relation

$$\hat{\phi}' = \cos(\phi - \phi') \hat{\phi} + \sin(\phi - \phi') \hat{\rho} \quad (3.52)$$

We can safely approximate the denominator term in Eq. (3.51) as $R \approx r$ without appreciably changing the solution. The exponent term, however, must preserve the phase information and therefore we will use R exactly as defined by Eq. (3.50). Substituting Eq. (3.52) and Eq. (3.50) into Eq. (3.51) we get

$$\bar{A} = \frac{a\mu_0}{4\pi} \int_0^{2\pi} I_0 (\cos(\phi - \phi') \hat{\phi} + \sin(\phi - \phi') \hat{\rho}) \frac{e^{-jk(r - a \sin \theta \cos(\phi - \phi'))}}{r} d\phi' \quad (3.53)$$

Since the loop is symmetric, the same solution will result regardless of the ϕ value chosen. For simplification purposes, let us choose $\phi = 0$. Also, since the $\hat{\rho}$ term has odd symmetry, the $\hat{\rho}$ term integrates to zero. Thus, Eq. (3.53) thus simplifies to

$$\bar{A} = \frac{a\mu_0 I_0 \hat{\phi}}{4\pi} \frac{e^{-jkr}}{r} \int_0^{2\pi} \cos \phi' e^{jka \sin \theta \cos \phi'} d\phi' \quad (3.54)$$

Equation (3.54) is solvable in closed form and the integral solution can be found in Gradshteyn and Ryzhik [6]. The vector potential solution is given as

$$\bar{A} = \frac{ja\mu_0 I_0 \hat{\phi}}{2} \frac{e^{-jkr}}{r} J_1(ka \sin \theta) \quad (3.55)$$

where J_1 is the Bessel function of the first kind and order 1.

The electric and magnetic field intensities can be found from Eq. (3.26) and Eq. (3.27) to be

$$E_{\phi} = \frac{ak\eta I_0}{2} \frac{e^{-jkr}}{r} J_1(ka \sin \theta) \quad (3.56)$$

where $E_r \approx 0$ and $E_{\theta} \approx 0$.

$$H_{\theta} = -\frac{E_{\phi}}{\eta} = -\frac{akI_0}{2} \frac{e^{-jkr}}{r} J_1(ka \sin \theta) \quad (3.57)$$

where H_r and $H_{\phi} \approx 0$.

Power density and radiation intensity. The power density and radiation intensity can be easily found to be

$$W_r(\theta, \phi) = \frac{1}{2\eta} |E_{\phi}|^2 = \frac{\eta}{8} \left(\frac{2\pi a}{\lambda} \right)^2 \frac{|I_0|^2}{r^2} J_1^2(ka \sin \theta) \quad (3.58)$$

$$U(\theta, \phi) = \frac{\eta}{8} \left(\frac{2\pi a}{\lambda} \right)^2 |I_0|^2 J_1^2(ka \sin \theta) \quad (3.59)$$

The variable ka can also be written as $\frac{2\pi a}{\lambda} = \frac{C}{\lambda}$ where C is the circumference of the loop. Figure 3.21 shows plots of the normalized radiation intensity given in Eq. (3.59) for $\frac{C}{\lambda} = 0.5, 1.25$, and 3. The radiation intensity is superimposed on a rectangular coordinate system.

Directivity. The directivity of the constant phasor current loop is again given by Eq. (3.16). We can substitute Eq. (3.59) into Eq. (3.16) to yield

$$D(\theta) = \frac{4\pi J_1^2(ka \sin \theta)}{\int_0^{2\pi} \int_0^{\pi} J_1^2(ka \sin \theta) \sin \theta d\theta d\phi} = \frac{2J_1^2\left(\frac{C}{\lambda} \sin \theta\right)}{\int_0^{\pi} J_1^2\left(\frac{C}{\lambda} \sin \theta\right) \sin \theta d\theta} \quad (3.60)$$

where $C = 2\pi a = \text{loop circumference}$.

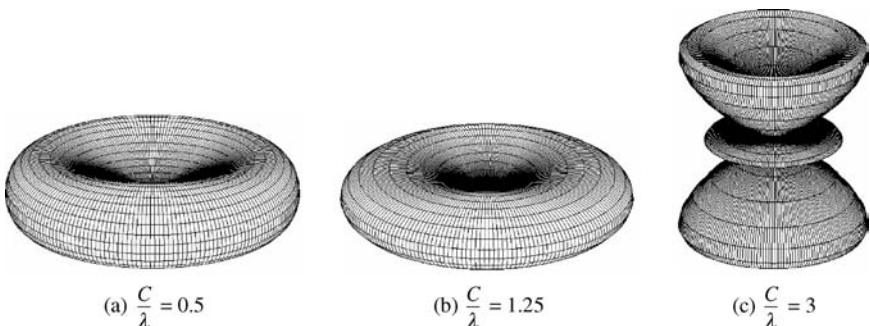


Figure 3.21 Loop radiation intensity.

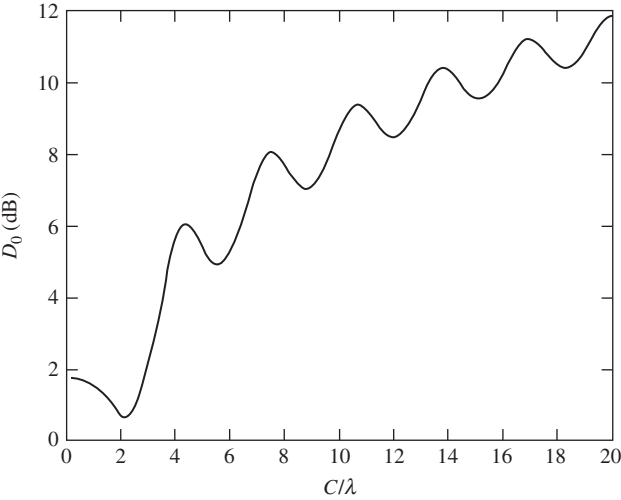


Figure 3.22 Maximum directivity of a circular loop.

The maximum directivity (D_0) is given when the numerator is maximum. The numerator is maximum when the Bessel function is maximum. For $C/\lambda > 1.84$ the Bessel function maximum is always 0.582. For $C/\lambda < 1.84$ the maximum is given as $2J_1^2(C/\lambda)$. We can plot the loop maximum directivity (D_0) vs. the circumference C in wavelengths to yield Fig. 3.22.

References

1. Balanis, C. *Antenna Theory Analysis and Design*, Wiley, 2d ed., New York, 1997.
2. Kraus, J., and R. Marhefka, *Antennas for All Applications*, 3d ed., McGraw-Hill, New York, 2002.
3. Stutzman, W.L., and G.A. Thiele, *Antenna Theory and Design*, Wiley, New York, 1981.
4. Haykin, S., *Communication Systems*, Wiley, page 723, New York, 2001.
5. King, R.W.P. "The Linear Antenna—Eighty Years of Progress," *Proceedings of the IEEE*, Vol. 55, pp. 2–16, Jan. 1967.
6. Gradshteyn, I.S., and I.M. Ryzhik, *Table of Integrals, Series, and Products*, Academic Press, New York, 1980.

Problems

3.1 For an antenna in free space creating an electric field intensity given by

$$\overline{E}_s = \frac{4 \cos \theta}{r} e^{-jkr} \hat{\theta} \text{ V/m}$$

- (a) What is the average power density?
- (b) What is the radiation intensity?
- (c) What is the total power radiated from the antenna?
- (d) Use MATLAB to plot the normalized power pattern.

62 Chapter Three

3.2 Use the help given in Example 3.4 and plot the following power patterns using MATLAB for $-90^\circ < \theta < 90^\circ$

- (a) $\cos^4(\theta)$
- (b) $\sin^2(\theta)$
- (c) $\sin^4(\theta)$

3.3 What is the 3-dB beamwidth for the patterns given in Prob. 3.2?

3.4 For the following radiation intensities, use MATLAB to plot the elevation plane polar pattern when $\phi = 0^\circ$ and the azimuthal plane polar pattern when $\theta = 90^\circ$. What are the boresight directions for these two patterns? Plot for the first pattern for $0^\circ < \theta < 180^\circ$ and the second pattern for $-90^\circ < \phi < 90^\circ$

- (a) $U(\theta, \phi) = \sin^2(\theta) \cos^2(\phi)$
- (b) $U(\theta, \phi) = \sin^6(\theta) \cos^6(\phi)$

3.5 Calculate the maximum directivity for the following radiation intensities

- (a) $U(\theta, \phi) = 4 \cos^2(\theta)$
- (b) $U(\theta, \phi) = 2 \sin^4(\theta)$
- (c) $U(\theta, \phi) = 2 \sin^2(\theta) \cos^2(\phi)$
- (d) $U(\theta, \phi) = 6 \sin^2(2\theta)$

3.6 Using Eq. (3.19) to find the beam solid angle for the following radiation intensities

- (a) $U(\theta, \phi) = 2 \cos^2(\theta)$
- (b) $U(\theta, \phi) = 4 \sin^2(\theta) \cos^2(\phi)$
- (c) $U(\theta, \phi) = 4 \cos^4(\theta)$

3.7 For the figure shown below calculate the receive power P_2 at antenna 2 if the transmit power $P_1 = 5$ kW. The transmitter gain is $G_1(\theta_1, \phi_1) = 4 \sin^4(\theta_1)$ and the receiver gain is $G_2(\theta_2, \phi_2) = 2 \sin^2(\theta_2)$. The operating frequency is 10 GHz. Use the Fig. P3.1 to solve the problem. The y - z and y' - z' planes are coplanar.

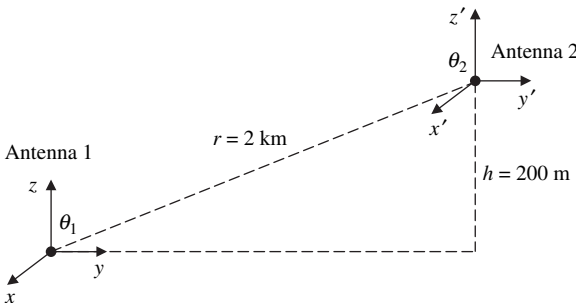


Figure P3.1 Two antennas separated by a distance r .

3.8 Create three normalized plots in one polar plot figure using MATLAB for the finite length dipole where $\frac{L}{\lambda} = 0.5, 1, \text{ and } 1.5$. Normalize each plot before plotting. One can use the “hold on” command to overlay the plots in the same figure.

3.9 What is the maximum directivity in dB for a finite length dipole whose length is $\frac{L}{\lambda} = 1.25$.

3.10 Create three normalized plots in one polar plot figure using MATLAB for the loop antenna where $\frac{C}{\lambda} = 0.5, 1.25, \text{ and } 3$. Normalize each plot before plotting. One can use the “hold on” command to overlay the plots in the same figure.

3.11 What is the maximum directivity in dB for the loop antenna whose circumference in wavelengths $\frac{C}{\lambda} = 10$.

This page intentionally left blank

Array Fundamentals

Smart antennas are composed of a collection of two or more antennas working in concert to establish a unique radiation pattern for the electromagnetic environment at hand. The antenna elements are allowed to work in concert by means of array element phasing, which is accomplished with hardware or is performed digitally. In the previous chapter we considered individual antenna elements such as the dipole or loop. In this chapter we will look at generic collections of antennas. These collections could be composed of dipole or loop antennas but it is not necessary to restrict ourselves to any particular antenna elements. As we will discover, the behavior of arrays transcends the specific elements used and the subject of arrays in and of itself has generated an extensive body of work. In fact the subject of arrays has merited entire textbooks devoted to the subject. Some very useful texts include Haykin [1], Johnson and Dudgeon [2], and Van Trees [3]. Extensive material can be found in the phased array text by Brookner [4] or in the article by Dudgeon [5].

Arrays of antennas can assume any geometric form. The various array geometries of common interest are linear arrays, circular arrays, planar arrays, and conformal arrays. We will begin with a discussion of linear arrays. A thorough treatment of linear arrays is found in the texts by Balanis [6] and Kraus and Marhefka [7].

4.1 Linear Arrays

The simplest array geometry is the linear array. Thus, all elements are aligned along a straight line and generally have a uniform interelement spacing. Linear arrays are the simplest to analyze and many valuable

insights can be gained by understanding their behavior. The minimum length linear array is the 2-element array.

4.1.1 Two element array

The most fundamental and simplest array to analyze is the two-element array. The two-element array demonstrates the same general behavior as much larger arrays and is a good starting point in order to understand the phase relationship between adjacent array elements. Figure 4.1 shows two vertically polarized infinitesimal dipoles aligned along the y axis and separated by a distance d . The field point is located at a distance r from the origin such that $r \gg d$. We can therefore assume that the distance vectors $\bar{r}_1$, $\bar{r}$, and $\bar{r}_2$ are all approximately parallel to each other.

We can therefore make the following approximations:

$$r_1 \approx r + \frac{d}{2} \sin \theta \quad (4.1)$$

$$r_2 \approx r - \frac{d}{2} \sin \theta \quad (4.2)$$

Let us additionally assume that the electrical phase of element 1 is $-\delta/2$ such that the phasor current in element 1 is $I_0 e^{-j\frac{\delta}{2}}$. The electrical phase of element 2 is $+\delta/2$ such that the phasor current in element 2 is $I_0 e^{j\frac{\delta}{2}}$. We can now find the distant electric field by using superposition as applied to these two dipole elements. Using Eq. (3.36) and Eqs. (4.1) and (4.2), and assuming that $r_1 \approx r_2 \approx r$ in the denominator, we can now find the total electric field.

$$\begin{aligned} E_\theta &= \frac{jk\eta I_0 e^{-j\frac{\delta}{2}} L \sin \theta}{4\pi r_1} e^{-jkr_1} + \frac{jk\eta I_0 e^{j\frac{\delta}{2}} L \sin \theta}{4\pi r_2} e^{-jkr_2} \\ &= \frac{jk\eta I_0 L \sin \theta}{4\pi r} e^{-jkr} \left[e^{-j\frac{(kd \sin \theta + \delta)}{2}} + e^{j\frac{(kd \sin \theta + \delta)}{2}} \right] \end{aligned} \quad (4.3)$$

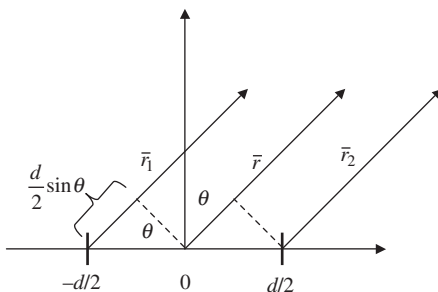


Figure 4.1 Two infinitesimal dipoles.

where δ = electrical phase difference between the two adjacent elements

L = dipole length

θ = angle as measured from the z axis in spherical coordinates

d = element spacing

We can further simplify Eq. (4.3) such that

$$E_{\theta} = \underbrace{\frac{jk\eta I_0 L e^{-jkr}}{4\pi r} \sin \theta}_{\text{Element factor}} \cdot \underbrace{\left(2 \cos \left(\frac{(kd \sin \theta + \delta)}{2} \right) \right)}_{\text{Array factor}} \quad (4.4)$$

where the element factor is the far field equation for one dipole and the array factor is the pattern function associated with the array geometry.

The distant field from an array of identical elements can always be broken down into the product of the element factor (EF) and the array factor (AF). The very fact that the antenna pattern can be multiplied by the array factor pattern demonstrates a property called *pattern multiplication*. Thus, the far field pattern of any array of antennas is always given by $(EF) \times (AF)$. The AF is dependent on the geometric arrangement of the array elements, the spacing of the elements, and the electrical phase of each element.

The normalized radiation intensity can be found by substituting Eq. (4.4) into Eq. (3.13) to get

$$\begin{aligned} U_n(\theta) &= [\sin \theta]^2 \cdot \left[\cos \left(\frac{(kd \sin \theta + \delta)}{2} \right) \right]^2 \\ &= [\sin \theta]^2 \cdot \left[\cos \left(\frac{\pi d}{\lambda} \sin \theta + \frac{\delta}{2} \right) \right]^2 \end{aligned} \quad (4.5)$$

We can demonstrate pattern multiplication by plotting Eq. (4.5) for the case where $d/\lambda = .5$ and $\delta = 0$. This is shown in Fig. 4.2. Figure 4.2a shows the power pattern for the dipole element alone. Figure 4.2b shows the array factor power pattern alone. And Fig. 4.2c shows the multiplication of the two patterns.

The overriding principle demonstrated by the 2-element array is that we can separate the element factor from the array factor. The array factor can be calculated for any array regardless of the individual elements chosen as long as all elements are the same. Thus, it is easier to first analyze arrays of isotropic elements. When the general array design is complete, one can implement the design by inserting the specific antenna elements required. Those antenna elements can include, but are not restricted to, dipoles, loops, horns, waveguide apertures, and patch antennas. A more exact representation of array radiation must

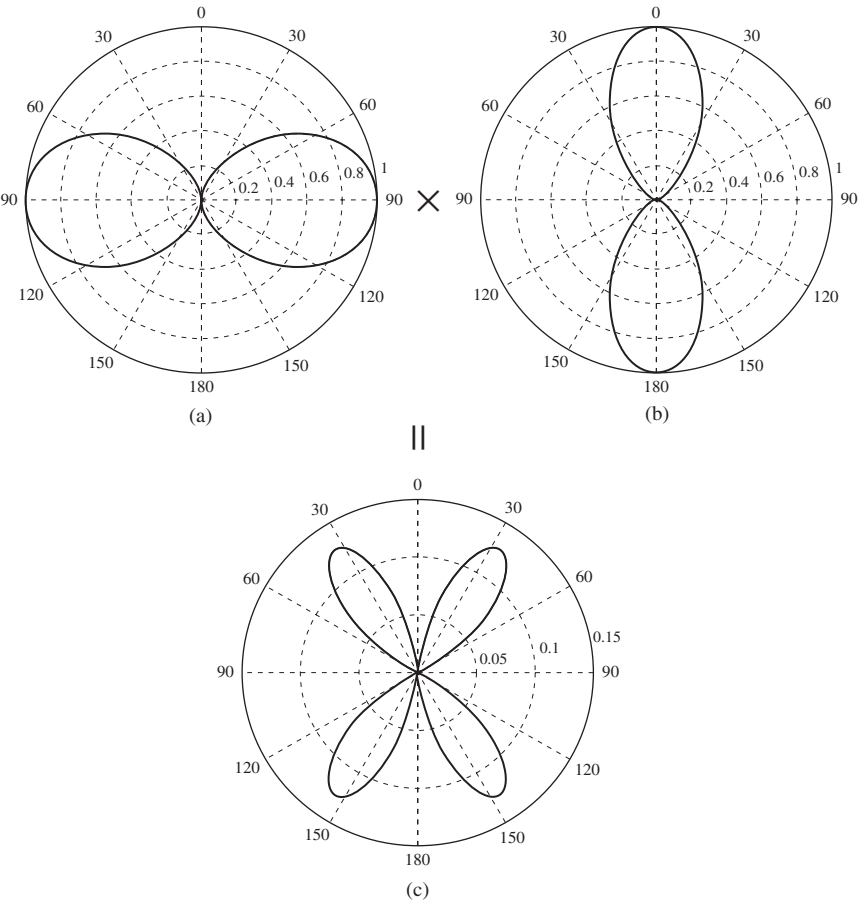


Figure 4.2 (a) Dipole pattern, (b) Array factor pattern, (c) Total pattern.

always include the effects of coupling between adjacent antenna elements. However, that topic is beyond the scope of this book and will be left for treatment by more advanced texts. The reader may refer to the text by Balanis [6] for more information on element coupling.

4.1.2 Uniform N -element linear array

The more general linear array is the N -element array. For simplification purposes, we will assume that all elements are equally spaced and have equal amplitudes. Later we may allow the antenna elements to have any arbitrary amplitude. Figure 4.3 shows an N -element linear array composed of isotropic radiating antenna elements. It is assumed that the n th element leads the $(n - 1)$ element by an electrical phase shift of δ radians. This phase shift can easily be implemented by shifting the phase of the antenna current for each element.

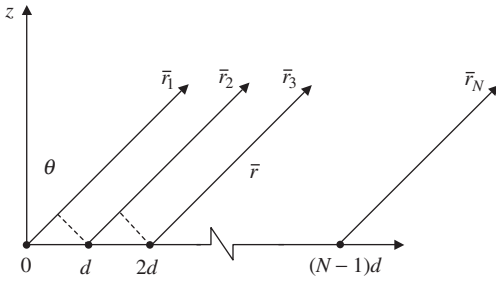


Figure 4.3 N -element linear array.

Assuming far field conditions such that $r \gg d$, we can derive the array factor as follows:

$$AF = 1 + e^{j(kd \sin \theta + \delta)} + e^{j2(kd \sin \theta + \delta)} + \dots + e^{j(N-1)(kd \sin \theta + \delta)} \quad (4.6)$$

where δ is the phase shift from element to element.

This series can more concisely be expressed by

$$AF = \sum_{n=1}^N e^{j(n-1)(kd \sin \theta + \delta)} = \sum_{n=1}^N e^{j(n-1)\psi} \quad (4.7)$$

where $\psi = kd \sin \theta + \delta$.

It should be noted that if the array is aligned along the z -axis then $\psi = kd \cos \theta + \delta$.

Since each isotropic element has unity amplitude, the entire behavior of this array is dictated by the phase relationship between the elements. The phase is directly proportional to the element spacing in wavelengths.

The array processing and array beamforming textbooks have taken an alternative approach to expressing Eq. (4.7). Let us begin by defining the array vector.

$$\tilde{a}(\theta) = \begin{bmatrix} 1 \\ e^{j(kd \sin \theta + \delta)} \\ \vdots \\ e^{j(N-1)(kd \sin \theta + \delta)} \end{bmatrix} = [1 \quad e^{j(kd \sin \theta + \delta)} \quad \dots \quad e^{j(N-1)(kd \sin \theta + \delta)}]^T \quad (4.8)$$

where $[]^T$ signifies the transpose of the vector within the brackets.

The vector $\tilde{a}(\theta)$ is a Vandermonde vector because it is in the form $[1 \ z \ \dots \ z^{(N-1)}]$. In the literature the array vector has been alternatively called: the *array steering vector* [2], the *array propagation vector* [8, 9], the *array response vector* [10], and the *array manifold vector* [3]. For simplicity's sake, we will call $\tilde{a}(\theta)$ the array vector. Therefore, the array

factor, in Eq. (4.7), can alternatively be expressed as the sum of the elements of the array vector.

$$\mathbf{AF} = \text{sum}(\tilde{\mathbf{a}}(\theta)) \quad (4.9)$$

The utility of the vector notation in Eq. (4.8) will be more readily seen in Chaps. 7 and 8 when we study angle-of-arrival estimation and smart antennas. In our current development, it is sufficient to use the notation of Eq. (4.7).

We may simplify the expression in Eq. (4.6) by multiplying both sides by $e^{j\psi}$ such that

$$e^{j\psi} \mathbf{AF} = e^{j\psi} + e^{j2\psi} + \dots + e^{jN\psi} \quad (4.10)$$

Subtracting Eq. (4.6) from Eq. (4.10) yields

$$(e^{j\psi} - 1)\mathbf{AF} = (e^{jN\psi} - 1) \quad (4.11)$$

The array factor can now be rewritten.

$$\begin{aligned} \mathbf{AF} &= \frac{(e^{jN\psi} - 1)}{(e^{j\psi} - 1)} = \frac{e^{j\frac{N}{2}\psi} (e^{j\frac{N}{2}\psi} - e^{-j\frac{N}{2}\psi})}{e^{j\frac{\psi}{2}} (e^{j\frac{\psi}{2}} - e^{-j\frac{\psi}{2}})} \\ &= e^{j\frac{(N-1)}{2}\psi} \frac{\sin(\frac{N}{2}\psi)}{\sin(\frac{\psi}{2})} \end{aligned} \quad (4.12)$$

The $e^{j\frac{(N-1)}{2}\psi}$ term accounts for the fact that the physical center of the array is located at $(N-1)d/2$. This array center produces a phase shift of $(N-1)\psi/2$ in the array factor. If the array is centered about the origin, the physical center is at 0 and Eq. (4.12) can be simplified to become

$$\mathbf{AF} = \frac{\sin(\frac{N}{2}\psi)}{\sin(\frac{\psi}{2})} \quad (4.13)$$

The maximum value of AF is when the argument $\psi = 0$. In that case $\mathbf{AF} = N$. This is intuitively obvious since an array of N elements should have a gain of N over a single element. We may normalize the AF to be reexpressed as

$$\mathbf{AF}_n = \frac{1}{N} \frac{\sin(\frac{N}{2}\psi)}{\sin(\frac{\psi}{2})} \quad (4.14)$$

In the cases where the argument $\psi/2$ is very small, we can invoke the small argument approximation for the $\sin(\psi/2)$ term to yield an approximation

$$\mathbf{AF}_n \approx \frac{\sin(\frac{N}{2}\psi)}{\frac{N}{2}\psi} \quad (4.15)$$

It should be noted that the array factor of Eq. (4.15) takes the form of a $\sin(x)/x$ function. This is because the uniform array itself presents a finite sampled rectangular window through which to radiate or receive a signal. The spatial Fourier transform of a rectangular window will yield a $\sin(x)/x$ function. The Fourier transform relationship between an antenna array and its radiation pattern is explained in [6].

Let us now determine the array factor nulls, maxima, and the mainlobe beamwidth.

Nulls From Eq. (4.15), the array nulls occur when the numerator argument $N\psi/2 = \pm n\pi$. Thus, the array nulls are given when

$$\frac{N}{2}(kd \sin \theta_{\text{null}} + \delta) = \pm n\pi$$

or

$$\theta_{\text{null}} = \sin^{-1} \left(\frac{1}{kd} \left(\pm \frac{2n\pi}{N} - \delta \right) \right) \quad n = 1, 2, 3 \dots \quad (4.16)$$

Since the $\sin(\theta_{\text{null}}) \leq 1$, for real angles, the argument in Eq. (4.16) must be ≤ 1 . Thus, only a finite set of n values will satisfy the equality.

Example 4.1 Find all of the nulls for an $N = 4$ element array with $d = .5\lambda$ and $\delta = 0$.

Solution Substituting N , d , and δ into Eq. (4.16)

$$\theta_{\text{null}} = \sin^{-1} \left(\pm \frac{n}{2} \right)$$

Thus, $\theta_{\text{null}} = \pm 30^\circ, \pm 90^\circ$

Maxima The mainlobe maximum in Eq. (4.15) occurs when the denominator term $\psi/2 = 0$. Thus

$$\theta_{\text{max}} = -\sin^{-1} \left(\frac{\delta \lambda}{2\pi d} \right) \quad (4.17)$$

The sidelobe maxima occur approximately when the numerator is a maximum. This occurs when the numerator argument $N\psi/2 = \pm(2n+1)\pi/2$. Thus

$$\begin{aligned} \theta_s &= \sin^{-1} \left(\frac{1}{kd} \left(\pm \frac{(2n+1)\pi}{N} - \delta \right) \right) \\ &= \pm \frac{\pi}{2} + \cos^{-1} \left(\frac{1}{kd} \left(\pm \frac{(2n+1)\pi}{N} - \delta \right) \right) \end{aligned} \quad (4.18)$$

Example 4.2 Find the mainlobe maximum and the sidelobe maxima for the case where $N = 4$, $d = .5\lambda$, and $\delta = 0$.

Solution Using Eq. (4.17) the mainlobe maximum can be found to be $\theta_{\max} = 0$ or π . π is a valid solution because the array factor is symmetric about the $\theta = \pi/2$ plane. The sidelobe maxima can be found from Eq. (4.18) to be

$$\theta_s = \pm 48.59^\circ, \pm 131.4^\circ$$

Beamwidth The beamwidth of a linear array is determined by the angular distance between the half-power points of the mainlobe. The mainlobe maximum is given by Eq. (4.17). Figure 4.4 depicts a typical normalized array radiation pattern with beamwidth as indicated.

The two half power points (θ_+ and θ_-) are found when the normalized $AF_n = .707$ ($AF_n^2 = .5$). If we use the array approximation given in Eq. (4.15) we can simplify the calculation of the beamwidth.

It is well known that $\sin(x)/x = .707$ when $x = \pm 1.391$. Thus, the normalized array factor is at the half-power points when

$$\frac{N}{2}(kd \sin \theta_{\pm} + \delta) = \pm 1.391 \quad (4.19)$$

Rearranging to solve for $\theta_{\pm}$ we get

$$\theta_{\pm} = \sin^{-1} \left(\frac{1}{kd} \left(\frac{\pm 2.782}{N} - \delta \right) \right) \quad (4.20)$$

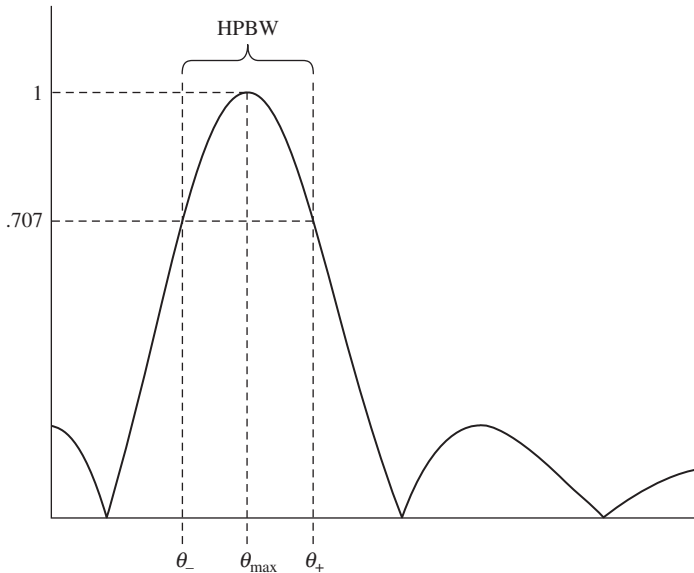


Figure 4.4 Half-power beamwidth of a linear array.

The half-power beamwidth is now easily shown to be

$$\text{HPBW} = |\theta_+ - \theta_-| \quad (4.21)$$

For large arrays, the beamwidth is narrow enough such that the HPBW can be approximated as

$$\text{HPBW} = 2|\theta_+ - \theta_{\max}| = 2|\theta_{\max} - \theta_-| \quad (4.22)$$

$\theta_{\max}$ is given by Eq. (4.17) and $\theta_{\pm}$ is given by Eq. (4.20).

Example 4.3 For the 4-element linear array with $\delta = -2.22$ and $d = .5\lambda$, what is the half-power beamwidth?

Solution $\theta_{\max}$ is first found using Eq. (4.17). Thus, $\theta_{\max} = 45^\circ$. θ_+ is found using Eq. (4.20). Thus, $\theta_+ = 68.13^\circ$. The half-power beamwidth is then approximated by using Eq. (4.21) to be $\text{HPBW} = 46.26^\circ$.

Broadside linear array. The most common mode of operation for a linear array is in the broadside mode. This is the case where $\delta = 0$ such that all element currents are in phase. Figure 4.5 shows three polar plots for a 4-element array for element distances $d/\lambda = .25, .5$, and $.75$.

This array is called a *broadside array* because the maximum radiation is broadside to the array geometry. Two major lobes are seen because the broadside array is symmetric about the $\theta = \pm\pi/2$ line. As the array element spacing increases, the array physically is longer, thereby decreasing the mainlobe width. The general rule for array radiation is that the mainlobe width is inversely proportional to the array length.

End-fire linear array. The name *end-fire* indicates that this array's maximum radiation is along the axis containing the array elements. Thus, maximum radiation is "out the end" of the array. This case is achieved when $\delta = -kd$. Figure 4.6 shows three polar plots for the end-fire 4-element array for the distances $d/\lambda = .25, .5$, and $.75$.

It should be noted that the mainlobe width for the ordinary end-fire case is much greater than the mainlobe width for the broadside case. Thus, ordinary end-fire arrays do not afford the same beamwidth efficiency as the broadside array. The beamwidth efficiency in this context is the beamwidth available relative to the overall array length.

An increased directivity end-fire array has been developed by Hansen-Woodyard [11] where the phase shift is modified such that $\delta = -(kd + \frac{\pi}{N})$. This provides a dramatic improvement in the beamwidth and directivity. Since the Hansen-Woodyard 1938 paper is not generally accessible, a detailed derivation can be found in [5].

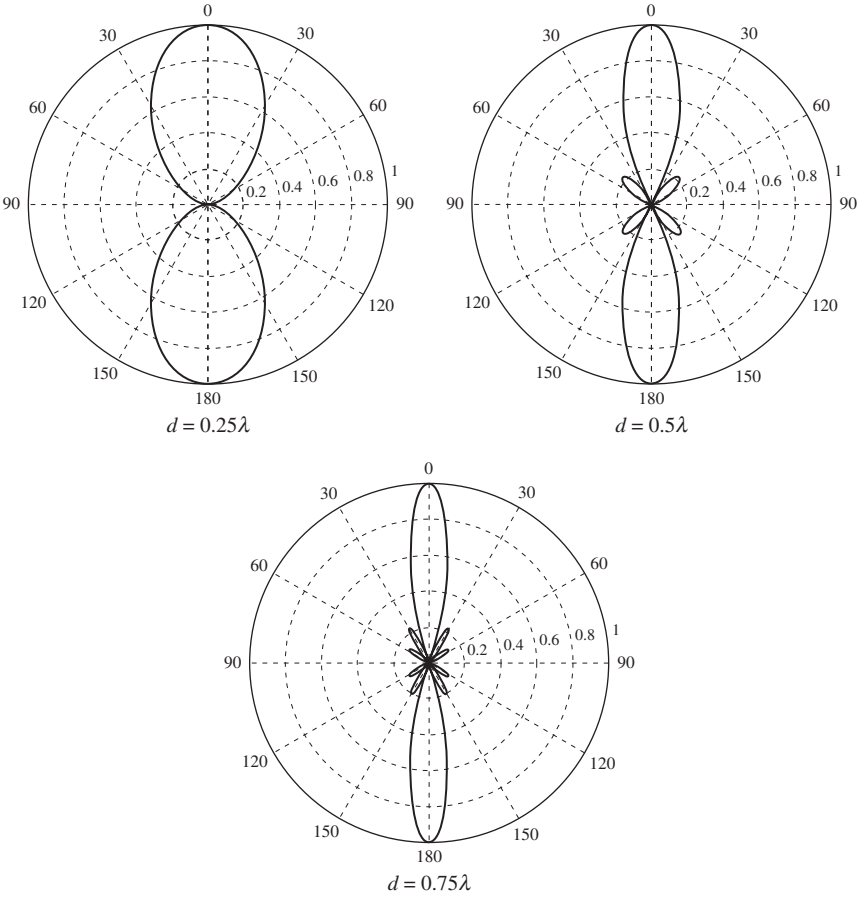


Figure 4.5 The 4-element broadside array with $\delta = 0$ and $d = .25\lambda$, $.5\lambda$, and $.75\lambda$.

Beamsteered linear array. A beamsteered linear array is an array where the phase shift δ is a variable thus allowing the mainlobe to be directed toward any direction of interest. The broadside and end-fire conditions are special cases of the more generalized beamsteered array. The beamsteering conditions can be satisfied by defining the phase shift $\delta = -kd \sin \theta_0$. We may rewrite the array factor in terms of beamsteering such that

$$AF_n = \frac{1}{N} \frac{\sin\left(\frac{Nkd}{2}(\sin \theta - \sin \theta_0)\right)}{\sin\left(\frac{kd}{2}(\sin \theta - \sin \theta_0)\right)} \quad (4.23)$$

Figure 4.7 shows polar plots for the beamsteered 8-element array for the $d/\lambda = .5$, and $\theta_0 = 20^\circ, 40^\circ$, and 60° . Major lobes exist above and below the horizontal because of array is symmetry.

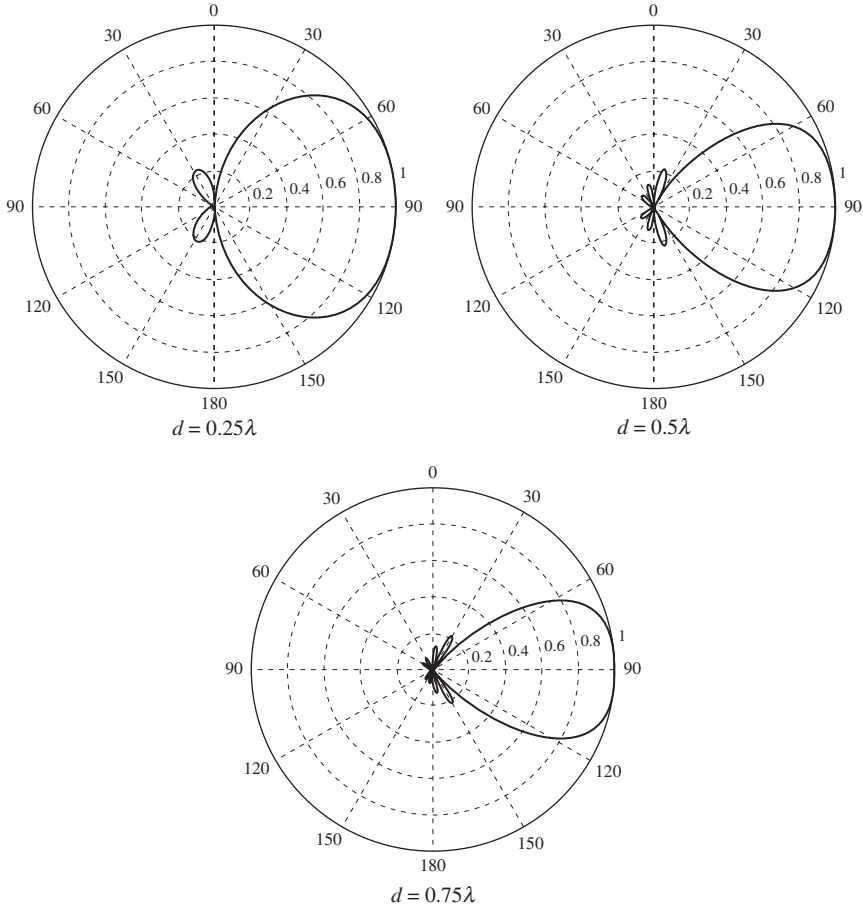


Figure 4.6 The 4-element end-fire array with $\delta = -kd$ and $d = .25\lambda$, $.5\lambda$, and $.75\lambda$.

The beamwidth of the beamsteered array can be determined by using Eqs. (4.20) and (4.21) such that

$$\theta_{\pm} = \sin^{-1} \left(\pm \frac{2.782}{Nkd} + \sin \theta_0 \right) \quad (4.24)$$

where $\delta = -kd \sin \theta_0$.

θ_0 = steering angle.

The beamsteered array beamwidth is now given as

$$\text{HPBW} = |\theta_+ - \theta_-| \quad (4.25)$$

For the $N = 6$ -element array, where $\theta_0 = 45^\circ$, $\theta_+ = 58.73^\circ$, and $\theta_- = 34.02^\circ$. Thus the beamwidth can be calculated to be $\text{HPBW} = 24.71^\circ$.

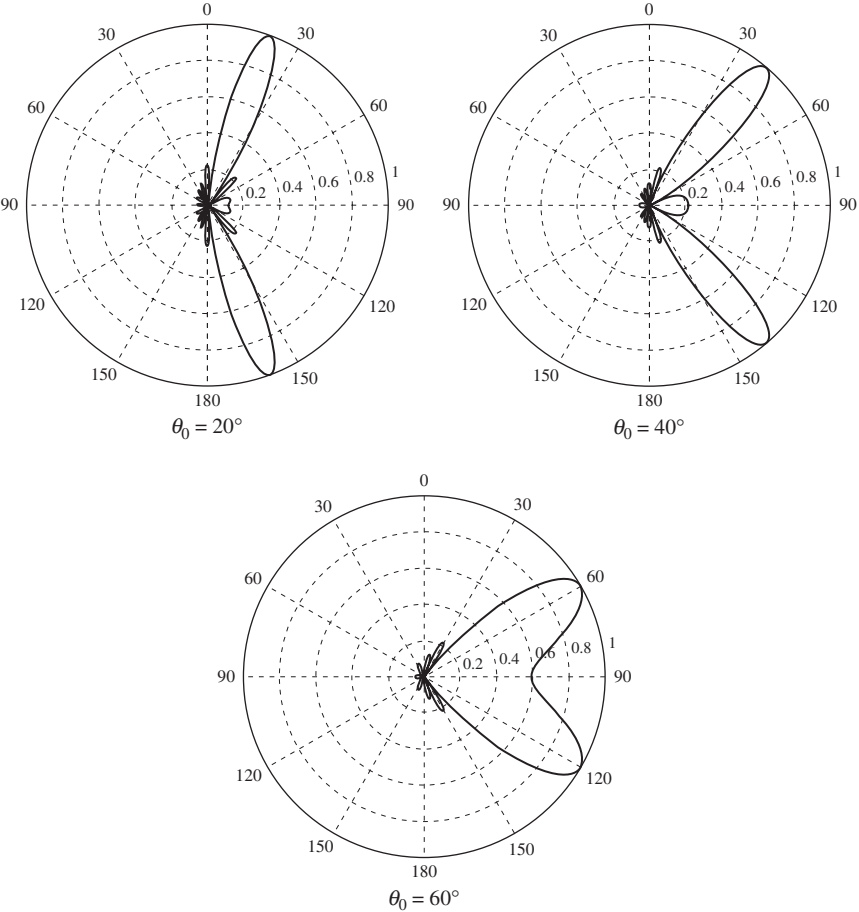


Figure 4.7 Beamsteered linear array with $\theta_0 = 45^\circ$.

4.1.3 Uniform N -element linear array directivity

Antenna directivity was previously defined in Eq. (3.16). Directivity is a measure of the antennas ability to preferentially direct energy in certain directions. The directivity equation is repeated as follows:

$$D(\theta, \phi) = \frac{4\pi U(\theta, \phi)}{\int_0^{2\pi} \int_0^\pi U(\theta, \phi) \sin \theta d\theta d\phi} \tag{4.26}$$

Our previous derivation for the array factor assumed the array was aligned along the horizontal axis. This derivation helped us to visualize the array performance relative to a broadside reference angle.

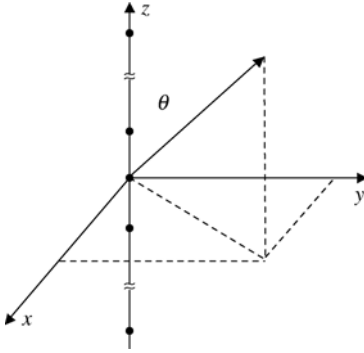


Figure 4.8 The N -element linear array along z -axis.

However, the horizontal array does not fit symmetrically into spherical coordinates. In order to simplify the calculation of the directivity, let us align the linear array along the z -axis as shown in Fig. 4.8.

Since we have now rotated the array by 90° to make it vertical, we can modify the AF by allowing $\psi = kd \cos \theta + \delta$. Now the broadside angle is when $\theta = 90^\circ$. Since the array factor is proportional to the signal level and not the power, we must square the array factor to yield the array radiation intensity $U(\theta)$. We now substitute the normalized approximate $(AF_n)^2$ into Eq. (4.26).

$$D(\theta) = \frac{4\pi \left(\frac{\sin(\frac{N}{2}(kd \cos \theta + \delta))}{\frac{N}{2}(kd \cos \theta + \delta)} \right)^2}{\int_0^{2\pi} \int_0^\pi \left(\frac{\sin(\frac{N}{2}(kd \cos \theta + \delta))}{\frac{N}{2}(kd \cos \theta + \delta)} \right)^2 \sin \theta \, d\theta \, d\phi} \quad (4.27)$$

The maximum value of the normalized array factor is unity. Thus the maximum directivity is given as

$$D_0 = \frac{4\pi}{\int_0^{2\pi} \int_0^\pi \left(\frac{\sin(\frac{N}{2}(kd \cos \theta + \delta))}{\frac{N}{2}(kd \cos \theta + \delta)} \right)^2 \sin \theta \, d\theta \, d\phi} \quad (4.28)$$

Solving for the maximum directivity is now simply the matter of solving the denominator integral although the integral itself is not trivial.

Broadside array maximum directivity. As was noted before, the case of broadside maximum directivity requires that $\delta = 0$. We can simplify the directivity equation by integrating over the ϕ variable. Thus, Eq. (4.28) can be simplified to be

$$D_0 = \frac{2}{\int_0^\pi \left(\frac{\sin(\frac{N}{2}kd \cos \theta)}{\frac{N}{2}kd \cos \theta} \right)^2 \sin \theta \, d\theta} \quad (4.29)$$

We can define the variable $x = \frac{N}{2}kd \cos \theta$. Then $dx = -\frac{N}{2}kd \sin \theta d\theta$. Substituting the new variable x into Eq. (4.29) yields

$$D_0 = \frac{Nkd}{\int_{-Nkd/2}^{Nkd/2} \left(\frac{\sin(x)}{x} \right)^2 dx} \quad (4.30)$$

As $Nkd/2 \gg \pi$, the limits can be extended to infinity without a significant loss in accuracy. The integral solution can be found in integral tables. Thus

$$D_0 \approx 2N \frac{d}{\lambda} \quad (4.31)$$

End-fire array maximum directivity. End-fire radiation conditions are achieved when the electrical phase between elements is $\delta = -kd$. This is equivalent to writing the overall phase term as $\psi = kd(\cos \theta - 1)$. Rewriting the maximum directivity, we get

$$D_0 = \frac{2}{\int_0^\pi \left(\frac{\sin(\frac{N}{2}kd(\cos \theta - 1))}{\frac{N}{2}kd(\cos \theta - 1)} \right)^2 \sin \theta d\theta} \quad (4.32)$$

We may again make a change of variable such that $x = \frac{N}{2}kd(\cos \theta - 1)$. Then $dx = -\frac{N}{2}kd \sin \theta d\theta$. Upon substitution of x into Eq. (4.32) we have

$$D_0 = \frac{Nkd}{\int_0^{Nkd} \left(\frac{\sin(x)}{x} \right)^2 dx} \quad (4.33)$$

As $Nkd/2 \gg \pi$, the upper limit can be extended to infinity without a significant loss in accuracy. Thus

$$D_0 \approx 4N \frac{d}{\lambda} \quad (4.34)$$

The end-fire array has twice the directivity of the broadside array. This is true because the end-fire array only has one major lobe, while the broadside array has two symmetric major lobes.

Beamsteered array maximum directivity. The most general case for the array directivity is found by defining the element-to-element phase shift δ in terms of the steering angle θ_0 . Let us rewrite Eq. (4.27) by substituting $\delta = -kd \cos \theta_0$.

$$D(\theta, \theta_0) = \frac{4\pi \left(\frac{\sin(\frac{N}{2}(kd(\cos \theta - \cos \theta_0)))}{\frac{N}{2}(kd(\cos \theta - \cos \theta_0))} \right)^2}{\int_0^{2\pi} \int_0^\pi \left(\frac{\sin(\frac{N}{2}(kd(\cos \theta - \cos \theta_0)))}{\frac{N}{2}(kd(\cos \theta - \cos \theta_0))} \right)^2 \sin \theta d\theta d\phi} \quad (4.35)$$

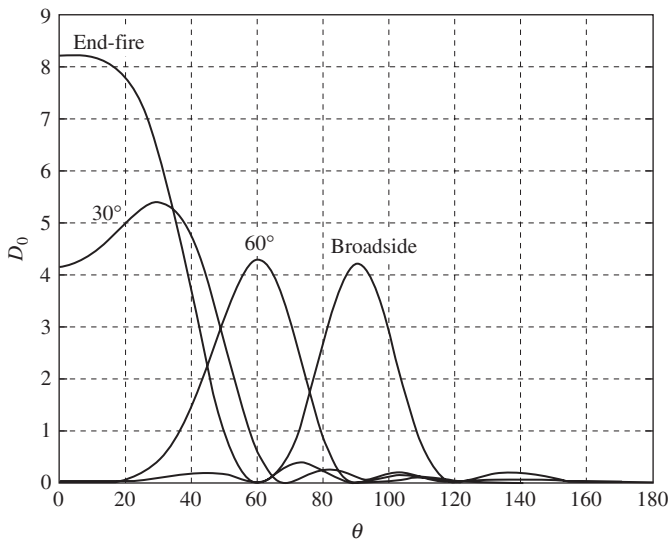


Figure 4.9 Family of steered array maximum directivity curves.

It is instructive to plot the linear array directivity vs. angle for a few different steering angles using MATLAB. We would expect the end-fire and broadside maximum directivities, calculated previously, to be two points on the more general plot of Eq. (4.35). Allowing $N = 4$ and $d = .5\lambda$, the directivity is shown plotted in Fig. 4.9. The maximum values are slightly higher than predicted by Eqs. (4.31) and (4.34) because no approximations were made to simplify the integration limits.

4.2 Array Weighting

The previous derivation for the array factor assumed that all of the isotropic elements had unity amplitude. Because of this assumption, the AF could be reduced to a simple series and a simple $\sin(x)/x$ approximation.

It was apparent from Figs. 4.5 and 4.6 that the array factor has sidelobes. For a uniformly weighted linear array, the largest sidelobes are down approximately 24 percent from the peak value. The presence of sidelobes means that the array is radiating energy in unintended directions. Additionally, due to reciprocity, the array is receiving energy from unintended directions. In a multipath environment, the sidelobes can receive the same signal from multiple angles. This is the basis for fading experienced in communications. If the direct transmission angle is known, it is best to steer the beam toward the desired direction and to

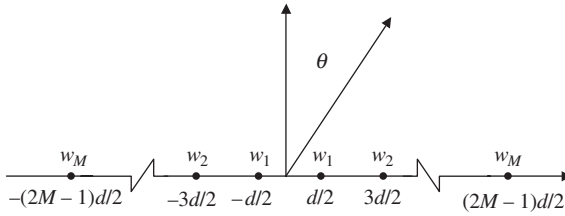


Figure 4.10 Even array with weights.

shape the sidelobes to suppress unwanted signals. The sidelobes can be suppressed by *weighting*, *shading*, or *windowing* the array elements. These terms are taken from the EM, underwater acoustics, and array signal processing communities, respectively. Array element weighting has numerous applications in areas such as digital signal processing (DSP), radio astronomy, radar, sonar, and communications. Two excellent foundational articles on array weighting are written by Harris [12] and Nuttall [13].

Figure 4.10 shows a symmetric linear array with an even number of elements N . The array is symmetrically weighted with weights as indicated:

The array factor is found by summing the weighted outputs of each element such that

$$\begin{aligned} \text{AF}_{\text{even}} = & w_M e^{-j \frac{(2M-1)}{2} k d \sin \theta} + \dots + w_1 e^{-j \frac{1}{2} k d \sin \theta} \\ & + w_1 e^{j \frac{1}{2} k d \sin \theta} + \dots + w_M e^{j \frac{(2M-1)}{2} k d \sin \theta} \end{aligned} \quad (4.36)$$

where $2M = N =$ total number of array elements. Each apposing pair of exponential terms in Eq. (4.36) forms complex conjugates. We can invoke the Euler's identity for the cosine to recast the even array factor given as follows:

$$\text{AF}_{\text{even}} = 2 \sum_{n=1}^M w_n \cos \left(\frac{(2n-1)}{2} k d \sin \theta \right) \quad (4.37)$$

Without loss of generality, the 2 can be eliminated from the expression in Eq. (4.37) to produce a quasi-normalization.

$$\text{AF}_{\text{even}} = \sum_{n=1}^M w_n \cos((2n-1)u) \quad (4.38)$$

where $u = \frac{\pi d}{\lambda} \sin \theta$.

The array factor is maximum when the argument is zero, implying $\theta = 0$. The maximum is then the sum of all of the array weights. Thus,

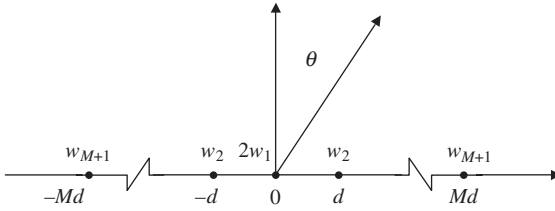


Figure 4.11 Odd array with weights.

we may completely normalize AF_{even} to be

$$AF_{\text{even}} = \frac{\sum_{n=1}^M w_n \cos((2n-1)u)}{\sum_{n=1}^M w_n} \quad (4.39)$$

It is easiest to express the array factor in the form of Eq. (4.33). However, for plotting purposes, it is best to use the normalized array factor given in Eq. (4.39).

An odd array is depicted in Fig. 4.11 with the center element at the origin.

We may again sum all of the exponential contributions from each array element to get the quasi-normalized odd array factor.

$$AF_{\text{odd}} = \sum_{n=1}^{M+1} w_n \cos(2(n-1)u) \quad (4.40)$$

where $2M+1 = N$.

In order to normalize Eq. (4.40) we must again divide by the sum of the array weights to get

$$AF_{\text{odd}} = \frac{\sum_{n=1}^{M+1} w_n \cos(2(n-1)u)}{\sum_{n=1}^{M+1} w_n} \quad (4.41)$$

We may alternatively express Eqs. (4.38) and (4.40) using the array vector nomenclature previously addressed in Eq. (4.8). Then, the array factor can be expressed in vector terms as

$$AF = \bar{w}^T \cdot \bar{a}(\theta) \quad (4.42)$$

where $\bar{a}(\theta)$ = array vector

$$\bar{w}^T = [w_M \quad w_{M-1} \quad \dots \quad w_1 \quad \dots \quad w_{M-1} \quad w_M]$$

The weights w_n can be chosen to meet any specific criteria. Generally the criterion is to minimize the sidelobes or possibly to place nulls at certain angles. However, symmetric scalar weights can only be utilized to shape sidelobes.

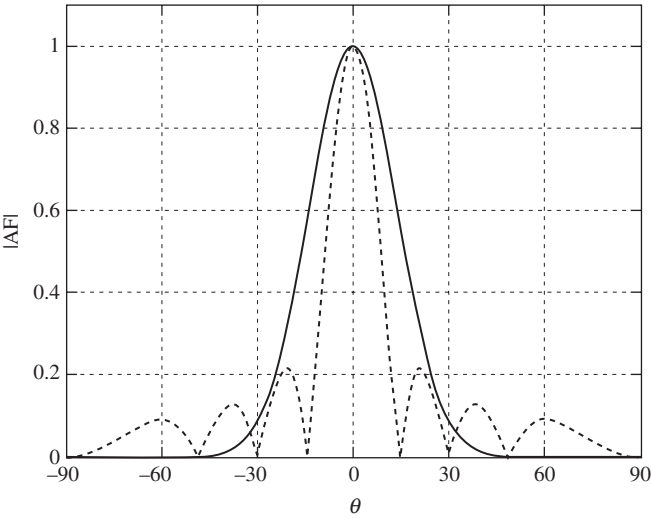


Figure 4.13 Array factor with binomial weights.

plotted in Fig. 4.12 using the stem command. The weighted array factor is superimposed over the unweighted array factor as shown in Fig. 4.13. The price paid for suppressing the sidelobes is seen in the broadening of the main beamwidth.

Blackman The Blackman weights are defined by

$$w(k+1) = .42 - .5 \cos(2\pi k/(N-1)) + .08 \cos(4\pi k/(N-1)) \quad k = 0, 1, \dots, N-1 \quad (4.43)$$

For the $N = 8$ element array, the normalized Blackman weights are $w_1 = 1$, $w_2 = .4989$, $w_3 = .0983$, and $w_4 = 0$. The eight array weights can be found using the Blackman(N) command in MATLAB. The normalized array weights are shown plotted in Fig. 4.14 using the stem command. The weighted array factor is shown in Fig. 4.15.

Hamming The Hamming weights are given by

$$w(k+1) = .54 - .46 \cos[2\pi k/(N-1)] \quad k = 0, 1, \dots, N-1 \quad (4.44)$$

The normalized Hamming weights are $w_1 = 1$, $w_2 = .673$, $w_3 = .2653$, and $w_4 = .0838$. The eight array weights can be found using the hamming(N) command in MATLAB. The normalized array weights are shown plotted in Fig. 4.16 using the stem command. The weighted array factor is shown in Fig. 4.17.

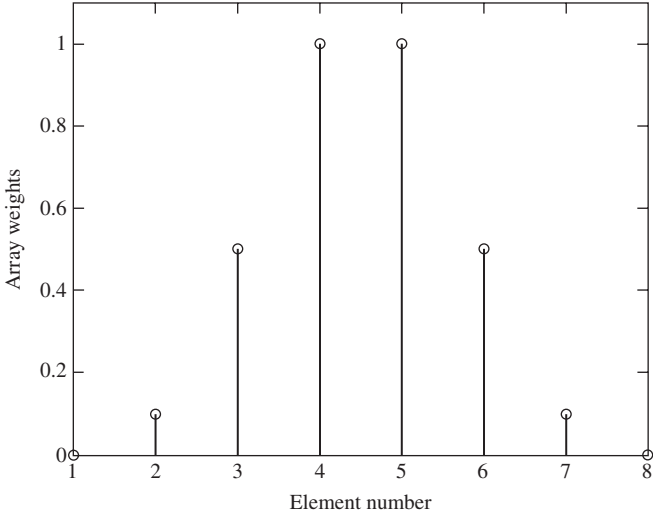


Figure 4.14 Blackman array weights.

Gaussian The gaussian weights are determined by the gaussian function to be

$$W(k+1) = e^{-\frac{1}{2} \left(\alpha \frac{k-N/2}{N/2} \right)^2} \quad k = 0, 1, \dots, N \quad \alpha \geq 2 \quad (4.45)$$

The normalized Gaussian weights for $\alpha = 2.5$ are $w_1 = 1$, $w_2 = .6766$, $w_3 = .3098$, and $w_4 = .0960$. The eight array weights can be found using the `gausswin(N)` command in MATLAB. The normalized array weights

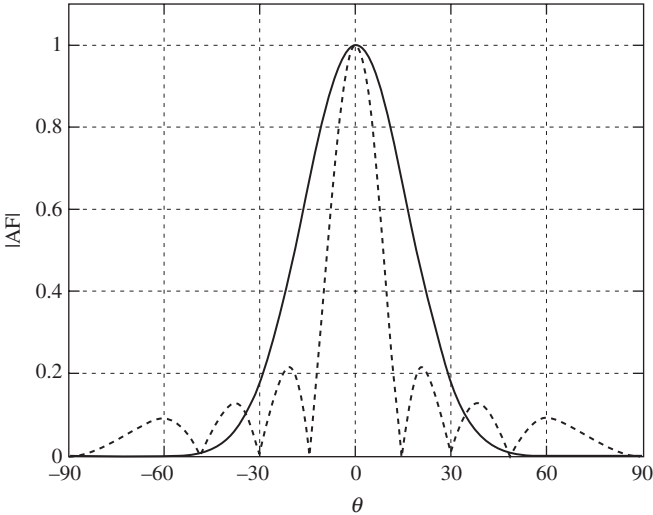


Figure 4.15 Array factor with Blackman weights.

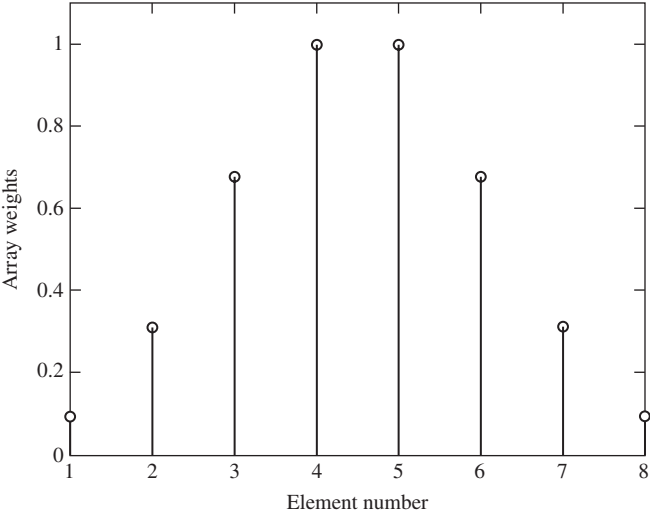


Figure 4.16 Hamming array weights.

are shown plotted in Fig. 4.18 using the stem command. The weighted array factor is shown in Fig. 4.19.

Kaiser-Bessel The Kaiser-Bessel weights are determined by

$$w(k) = \frac{I_0 \left[\pi \alpha \sqrt{1 - \left(\frac{k}{N/2} \right)^2} \right]}{I_0 [\pi \alpha]} \quad k = 0, 1, \dots, N/2 \quad \alpha > 1 \quad (4.46)$$

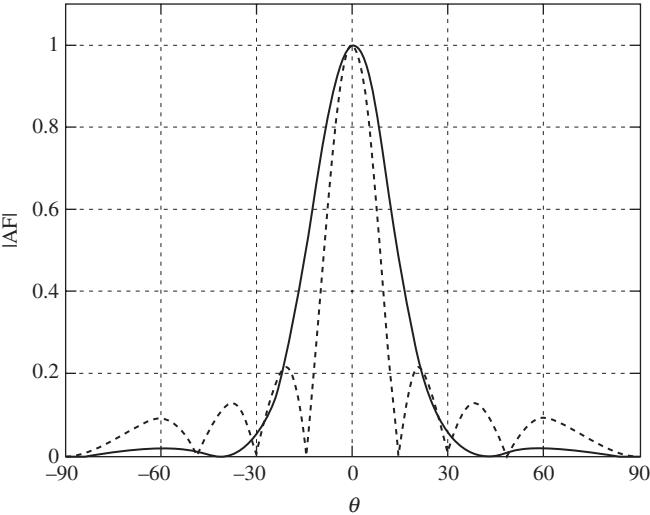


Figure 4.17 Array factor with Hamming weights.

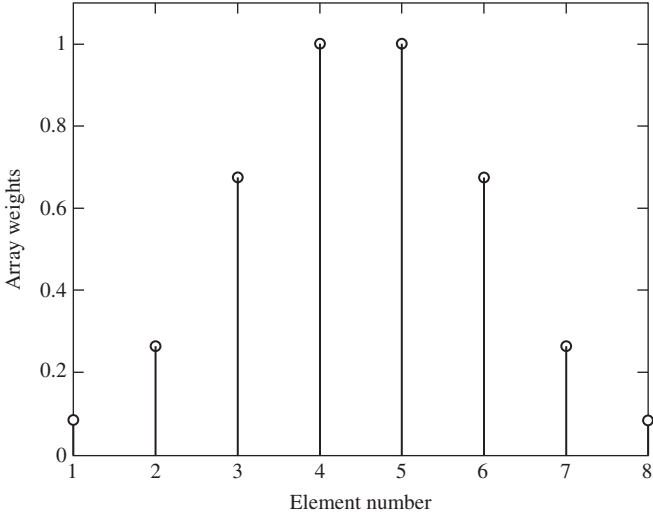


Figure 4.18 Gaussian array weights.

The normalized Kaiser-Bessel weights for $\alpha = 3$ are $w_1 = 1, w_2 = .8136, w_3 = .5137$, and $w_4 = .210$. The eight array weights can be found using the `kaiser(N,  $\alpha$ )` command in MATLAB. The normalized array weights are shown plotted in Fig. 4.20 using the `stem` command. The weighted array factor is shown in Fig. 4.21.

It should be noted that the Kaiser-Bessel weights provide one of the lowest array sidelobe levels while still maintaining nearly the same

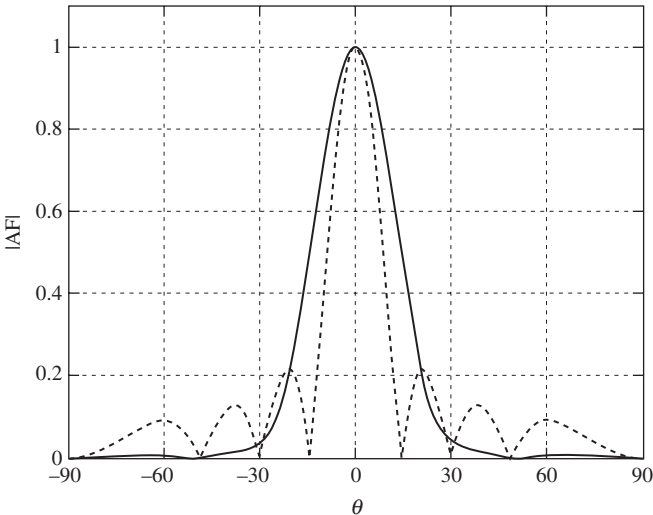


Figure 4.19 Array factor with gaussian weights.

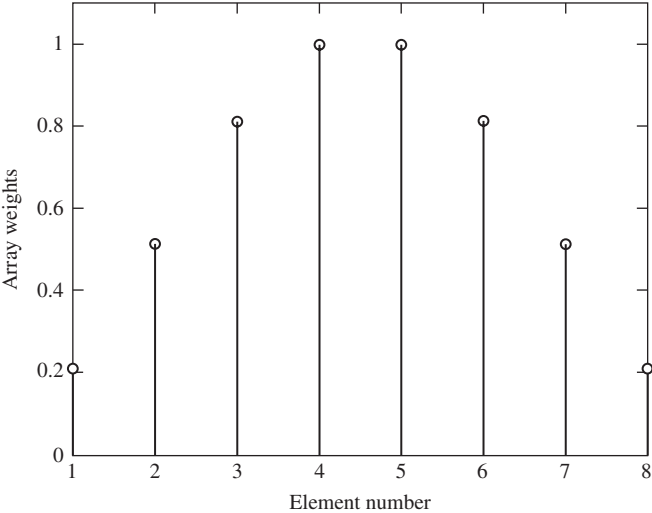


Figure 4.20 Kaiser-Bessel array weights.

beamwidth as uniform weights. Additionally, the Kaiser-Bessel weights for $\alpha = 1$ yield the uniform set of weights.

Other potential weight functions are the Blackman-Harris, Bohman, Hanning, Bartlett, Dolph-Chebyshev, and Nuttall. A detailed description of these functions can be found in [12, 13]. In addition, many of these weight functions are available in MATLAB.

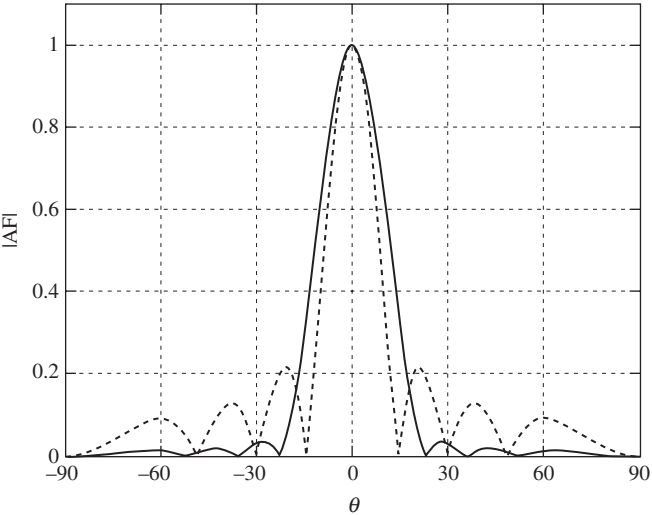


Figure 4.21 Array factor with Kaiser-Bessel weights.

4.2.1 Beamsteered and weighted arrays

Previously, in Sec. 4.1.3 we discussed the beamsteered uniformly weighted array. We could steer the mainlobe to any desired direction but we still experienced the problem of relatively large minor lobes. The nonuniformly weighted array can also be modified in order to steer the beam to any direction desired and with suppressed sidelobe levels.

We can repeat Eqs. (4.38) and (4.40) but we can modify them to include beamsteering.

$$AF_{\text{even}} = \sum_{n=1}^M w_n \cos((2n-1)u) \quad (4.47)$$

$$AF_{\text{odd}} = \sum_{n=1}^{M+1} w_n \cos(2(n-1)u) \quad (4.48)$$

where

$$u = \frac{\pi d}{\lambda} (\sin \theta - \sin \theta_0)$$

As an example, we can use Kaiser-Bessel weights and steer the mainlobe to three separate angles. Let $N = 8$, $d = \lambda/2$, $\alpha = 3$, $w_1 = 1$, $w_2 = .8136$, $w_3 = .5137$, and $w_4 = .210$. The beamsteered array factor for the weighted even element array is shown in Fig. 4.22.

In general, any array can be steered to any direction by either using phase shifters in the hardware or by digitally phase shifting the data at

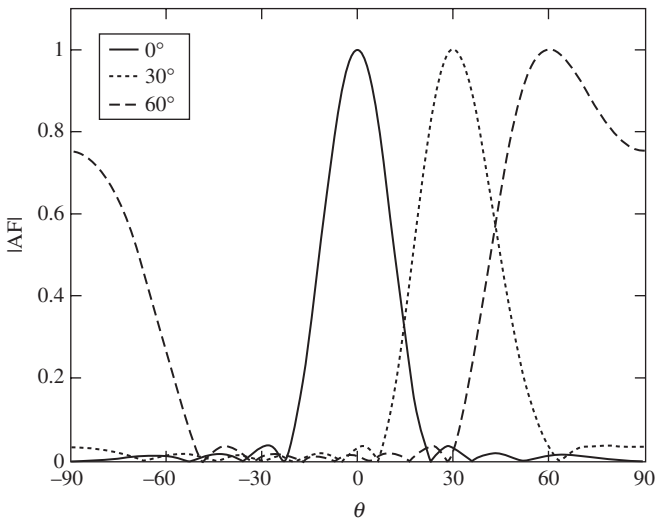


Figure 4.22 Beamsteered Kaiser-Bessel weighted array.

the back end of the receiver. If the received signal is digitized and processed, this signal processing is often called *digital beamforming* (DBF) [8]. Current technologies are making it more feasible to perform DBF and therefore allow the array designer to bypass the need for hardware phase shifters. The DBF performed can be used to steer the antenna beam according to any criteria specified by the user.

4.3 Circular Arrays

Linear arrays are very useful and instructive but there are occasions where a linear array is not appropriate for the building, structure, or vehicle upon which it is mounted. Other array geometries may be necessary to appropriately fit into a given scenario. Such additional arrays can include the circular array. Just as the linear array was used for increased gain and beamsteering, the circular array can also be used. Figure 4.23 shows a circular array of N elements in the x - y plane. The array has N elements and the array radius is a .

The n th array element is located at the radius a with the phase angle ϕ_n . Additionally, each element can have an associated weight w_n and phase δ_n . As before, with the linear array, we will assume far-field conditions and will assume that the observation point is such that the position vectors $\bar{r}$ and $\bar{r}_n$ are parallel. We can now define the unit vector in the direction of each array element n .

$$\hat{\rho}_n = \cos \phi_n \hat{x} + \sin \phi_n \hat{y}$$
 (4.49)

We can also define the unit vector in the direction of the field point.

$$\hat{r} = \sin \theta \cos \phi \hat{x} + \sin \theta \sin \phi \hat{y} + \cos \theta \hat{z}$$
 (4.50)

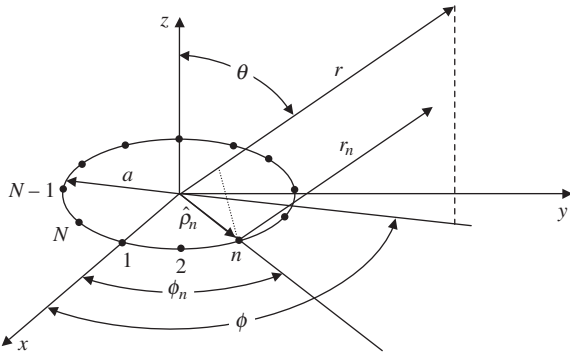


Figure 4.23 Circular array of N -elements.

It can be shown that the distance r_n is less than the distance r by the scalar projection of $\hat{\rho}_n$ onto $\hat{r}$. (This is indicated by the dotted line in Fig. 4.23). Thus,

$$r_n = r - a\hat{\rho}_n \cdot \hat{r} \quad (4.51)$$

with

$$\hat{\rho}_n \cdot \hat{r} = \sin \theta \cos \phi \cos \phi_n + \sin \theta \sin \phi \sin \phi_n = \sin \theta \cos(\phi - \phi_n)$$

The array factor can now be found in a similar fashion as was done with the linear array. With some effort, it can be shown that

$$\text{AF} = \sum_{n=1}^N w_n e^{-j(ka\hat{\rho}_n \cdot \hat{r} + \delta_n)} = \sum_{n=1}^N w_n e^{-j[ka \sin \theta \cos(\phi - \phi_n) + \delta_n]} \quad (4.52)$$

where

$$\phi_n = \frac{2\pi}{N}(n-1) = \text{angular location of each element.}$$

4.3.1 Beamsteered circular arrays

The beamsteering of circular arrays is identical in form to the beamsteering of linear arrays. If we beamsteer the circular array to the angles

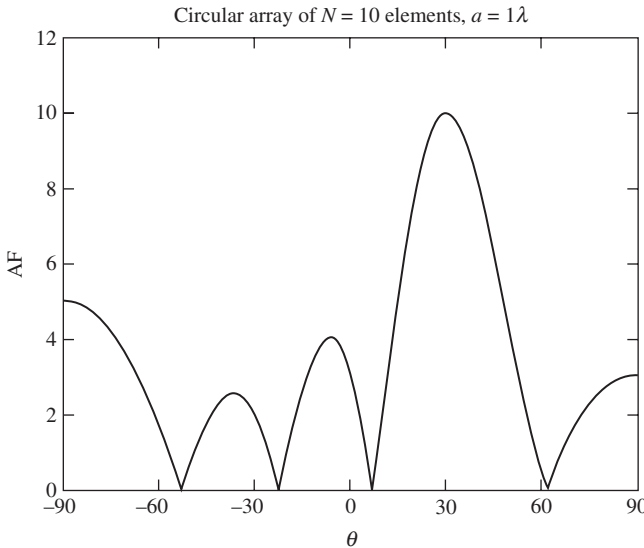


Figure 4.24 AF elevation pattern for beamsteered circular array ($\theta_0 = 30^\circ$, $\phi_0 = 0^\circ$).

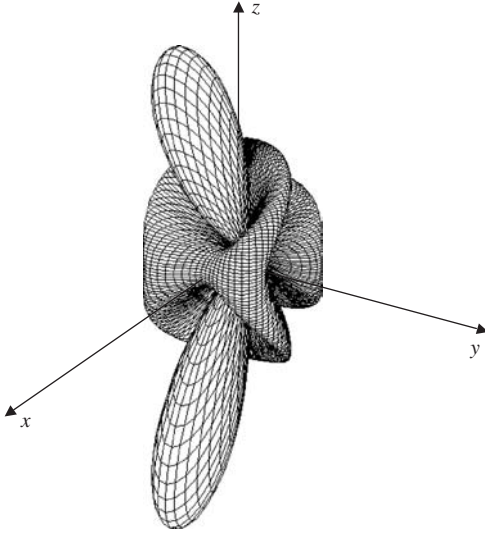


Figure 4.25 3-D AF pattern for beamsteered circular array ($\theta_0 = 30^\circ$, $\phi_0 = 0^\circ$).

(θ_0, ϕ_0) we can determine that the element to element phase angle is $\delta_n = -ka \sin \theta_0 \cos(\phi_0 - \phi_n)$. We can thus rewrite the array factor as

$$AF = \sum_{n=1}^N w_n e^{-j[ka[\sin \theta \cos(\phi - \phi_n) - \sin \theta_0 \cos(\phi_0 - \phi_n)]]} \quad (4.53)$$

The circular array AF can be plotted in two or three dimensions. Let us assume that all weights are uniform and that the array is steered to the angles $\theta_0 = 30^\circ$ and $\phi_0 = 0^\circ$. With $N = 10$ and $a = \lambda$, we can plot the elevation pattern in the $\phi = 0^\circ$ plane as shown in Fig. 4.24.

We can also plot the array factor in three dimensions as a mesh plot. Using the same parameters as above, we see the beamsteered circular array pattern in Fig. 4.25.

4.4 Rectangular Planar Arrays

Having explored the linear and circular arrays, we can move on to slightly more complex geometries by deriving the pattern for rectangular planar arrays. The following development is similar to that found in both Balanis [6] and in Johnson and Jasik [14].

Figure 4.26 shows a rectangular array in the x - y plane. There are M elements in the x -direction and N elements in the y -direction creating an $M \times N$ array of elements. The m -nth element has weight w_{mn} . The x -directed elements are spaced d_x apart and the y -directed elements are spaced d_y apart. The planar array can be viewed as M linear arrays of N elements or as N linear arrays of M elements. Since we already

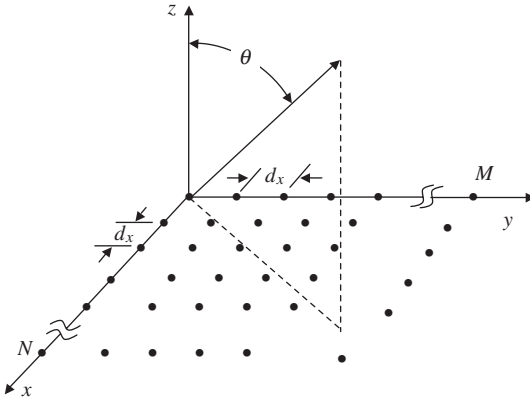


Figure 4.26 $N \times M$ rectangular planar array.

know the array factor for an M or N element array acting alone, we can use pattern multiplication to find the pattern of the entire $M \times N$ element array. Using pattern multiplication we have

$$\begin{aligned} \text{AF} &= \text{AF}_x \cdot \text{AF}_y = \sum_{m=1}^M a_m e^{j(m-1)(kd_x \sin \theta \cos \phi + \beta_x)} \sum_{n=1}^N b_n e^{j(n-1)(kd_y \sin \theta \sin \phi + \beta_y)} \\ &= \sum_{m=1}^M \sum_{n=1}^N w_{mn} e^{j[(m-1)(kd_x \sin \theta \cos \phi + \beta_x) + (n-1)(kd_y \sin \theta \sin \phi + \beta_y)]} \end{aligned} \quad (4.54)$$

where $w_{mn} = a_m \cdot b_n$

The weights a_m and b_n can be uniform or can be in any form according to the designer's needs. This could include the various weights discussed in Sec. 4.2 such as the binomial, Kaiser-Bessel, Hamming, or Gaussian weights. The a_m weights do not have to be identical to the b_n weights. Thus, we might choose the a_m weights to be binomial weights while the b_n weights are Gaussian. Any combination of weighting can be used and w_{mn} is merely the consequence of the product $a_m \cdot b_n$.

If beamsteering is desired then the phase delays β_x and β_y are given by

$$\beta_x = -kd_x \sin \theta_0 \cos \phi_0 \quad \beta_y = -kd_y \sin \theta_0 \sin \phi_0 \quad (4.55)$$

Example 4.4 Design and plot a 8×8 element array with equal element spacing such that $d_x = d_y = .5\lambda$. Let the array be beamsteered to $\theta_0 = 45^\circ$ and

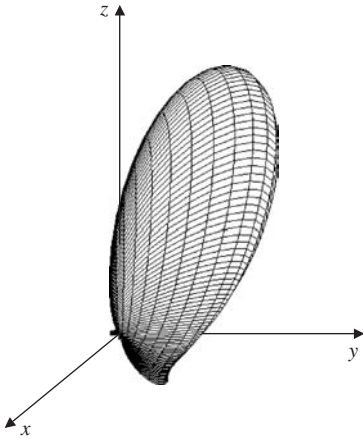


Figure 4.27 Beamsteered planar array pattern.

$\phi_0 = 45^\circ$. The element weights are chosen to be the Kaiser-Bessel weights given in Sec. 4.2. Plot the pattern for the range $0 \leq \theta \leq \pi/2$ and $0 \leq \phi \leq 2\pi$.

Solution Taking the weights from Sec. 4.2 we have $a_1 = a_4 = b_1 = b_4 = .2352$, and $a_2 = a_3 = b_2 = b_3 = 1$. Thus, $w_{1n} = w_{m1} = .0055$ and all other w_{mn} 's = 1. We can substitute the weights into Eq. (4.51) to calculate the array factor. MATLAB can be used to plot Fig. 4.27.

4.5 Fixed Beam Arrays

Fixed beam arrays are designed such that the array pattern consists of several simultaneous *spot* beams transmitting in fixed angular directions. Normally these directions are in equal angular increments so as to insure a relatively uniform coverage of a region in space. However, this is not a necessary restriction. These fixed beams can be used in satellite communications to create spot beams toward fixed earth-based locations. As an example, the *Iridium*¹ (low earth orbit) satellite constellation system has 48 spot beams per satellite. Spot beams are sometimes also called *pin-cushion* beams because of the similarity to pins in a pin cushion. Figure 4.28 shows an example of a planar array creating three spot beams.

Fixed beams can also be used for mobile communication base stations in order to provide space division multiple access (SDMA) capabilities. Several have treated the subject of fixed beam systems such as Mailloux [15], Hansen [16], and Pattan [17].

¹Iridium Satellite LLC

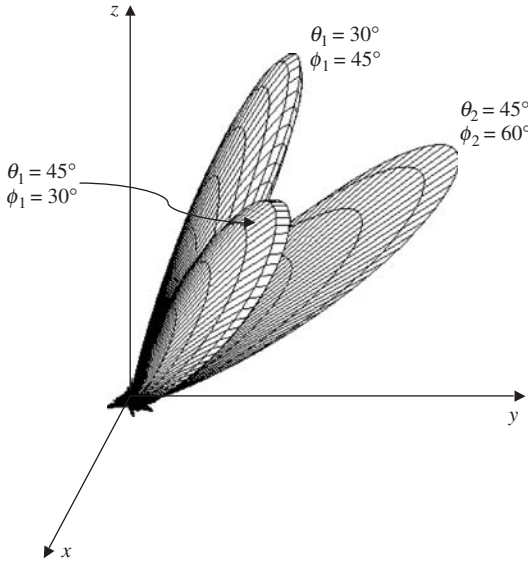


Figure 4.28 Three spot beams created by a 16×16 planar array.

4.5.1 Butler matrices

One method for easily creating fixed beams is through the use of Butler matrices. Details of the derivation can be found in Butler and Lowe [18], and Shelton and Kelleher [19]. The Butler matrix is an analog means of producing several simultaneous fixed beams through the use of phase shifters. As an example, let us assume a linear array of N elements. If $N = 2^n$ elements, the array factor can be given as

$$\text{AF}(\theta) = \frac{\sin(N\pi \frac{d}{\lambda} \sin \theta - \beta_\ell)}{N\pi \frac{d}{\lambda} \sin \theta - \beta_\ell} = \frac{\sin[N\pi \frac{d}{\lambda} (\sin \theta - \sin \theta_\ell)]}{N\pi \frac{d}{\lambda} (\sin \theta - \sin \theta_\ell)} \quad (4.56)$$

with

$$\begin{aligned} \sin \theta_\ell &= \frac{\ell \lambda}{Nd} \\ \beta_\ell &= \ell \pi \\ \ell &= \pm \frac{1}{2}, \pm \frac{3}{2}, \dots, \pm \frac{(N-1)}{2} \end{aligned}$$

The ℓ values create evenly spaced contiguous beams about $\theta = 0^\circ$. If the element spacing is $d = \lambda/2$, the beams are evenly distributed over the span of 180° . If $d > \lambda/2$, the beams span an ever decreasing range of angles. Grating lobes are not created provided that the only phase shifts used are those defined by β_ℓ .

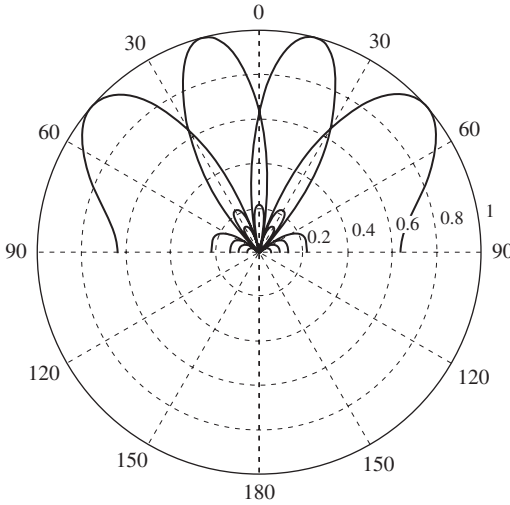


Figure 4.29 Scalloped beams using Butler approach.

As an example, let us choose an $N = 4$ element array with $d = \lambda/2$. We can use Eq. (4.56) to produce N fixed beams and plot the results using MATLAB. These beams are sometimes referred to as *scalloped* beams because of the resemblance to the scallop shell. Since $N = 4$, then $\ell = -\frac{3}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{3}{2}$. Substituting these values into the equation for $\sin \theta_\ell$, we get the polar plot as shown in Fig. 4.29.

These beams can be created using fixed phase shifters by noting that $\beta_\ell = \ell\pi = \pm\frac{\pi}{2}, \pm\frac{3\pi}{2}$. Figure 4.30 depicts a Butler matrix labyrinth of phase shifters for $N = 4$ -array elements.

As can be noted from the figure, the $1R$ port will yield the first beam to the right of broadside to the array. The $2R$ port yields the second beam to the right and so on. Thus, through the use of the Butler matrix labyrinth, one can simultaneously look in N directions with an N -element array.

4.6 Fixed Sidelobe Canceling

The basic goal of a fixed *sidelobe canceller* (SLC) is to choose array weights such that a null is placed in the direction of interference while the mainlobe maximum is in the direction of interest. The concept of an SLC was first presented by Howells in 1965 [20]. Since that time adaptive SLCs have been extensively studied. A detailed description for an adaptive SLC will be reserved for Chap. 8.

In our current development, we will discuss fixed sidelobe canceling for one fixed known desired source and two fixed undesired interferers. All signals are assumed to operate at the same carrier frequency. Let

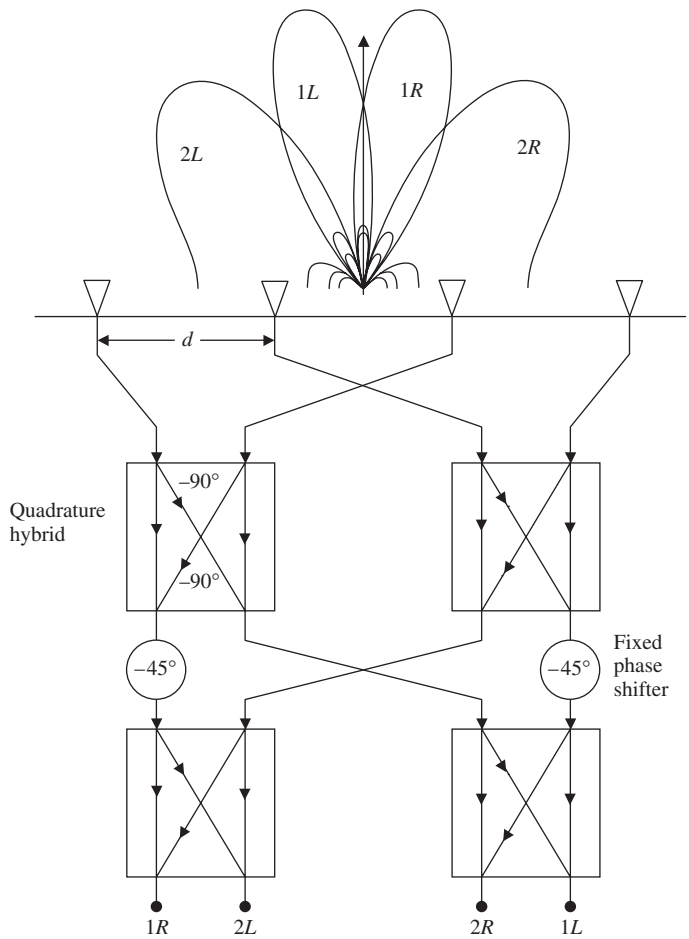


Figure 4.30 Butler matrix labyrinth for $N = 4$.

us assume a 3-element array with the desired signal and interferers as shown in Fig. 4.31.

The array vector is given by

$$\bar{a}(\theta) = [e^{-jkd \sin \theta} \quad 1 \quad e^{jkd \sin \theta}]^T \tag{4.57}$$

The, as yet to be determined, array weights are given by

$$\bar{w}^T = [w_1 \quad w_2 \quad w_3] \tag{4.58}$$

Therefore, the total array output from the summer is given as

$$S = \bar{w}^T \cdot \bar{a} = w_1 e^{-jkd \sin \theta} + w_2 + w_3 e^{jkd \sin \theta} \tag{4.59}$$

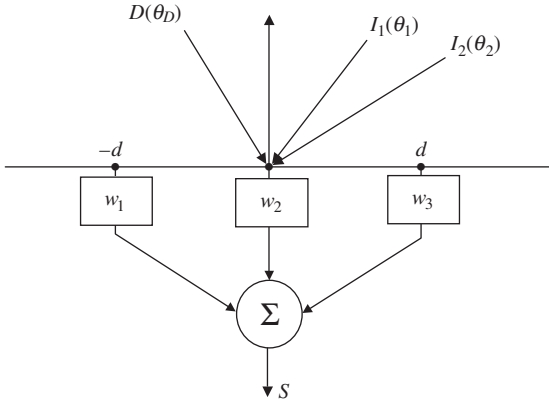


Figure 4.31 The 3-element array with desired and interfering signals.

The array output for the desired signal will be designated by S_D whereas the array output for the interfering signals will be designated by S_1 and S_2 . Since there are three unknown weights, there must be three conditions satisfied.

$$\text{Condition 1: } S_D = w_1 e^{-jkd \sin \theta_D} + w_2 + w_3 e^{jkd \sin \theta_D} = 1$$

$$\text{Condition 2: } S_1 = w_1 e^{-jkd \sin \theta_1} + w_2 + w_3 e^{jkd \sin \theta_1} = 0$$

$$\text{Condition 2: } S_2 = w_1 e^{-jkd \sin \theta_2} + w_2 + w_3 e^{jkd \sin \theta_2} = 0$$

Condition 1 demands that $S_D = 1$ for the desired signal, allowing the desired signal to be received without modification. Conditions 2 and 3 reject the undesired signals. These conditions can be recast in matrix form as

$$\begin{bmatrix} e^{-jkd \sin \theta_D} & 1 & e^{jkd \sin \theta_D} \\ e^{-jkd \sin \theta_1} & 1 & e^{jkd \sin \theta_1} \\ e^{-jkd \sin \theta_2} & 1 & e^{jkd \sin \theta_2} \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad (4.60)$$

One can invert the matrix to find the required complex weights w_1 , w_2 , and w_3 . As an example, if the desired signal is arriving from $\theta_D = 0^\circ$ while $\theta_1 = -45^\circ$ and $\theta_2 = 60^\circ$, the necessary weights can be calculated to be

$$\begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} 0.748 + 0.094i \\ -0.496 \\ 0.748 - 0.094i \end{bmatrix} \quad (4.61)$$

The array factor is shown plotted in Fig. 4.32.

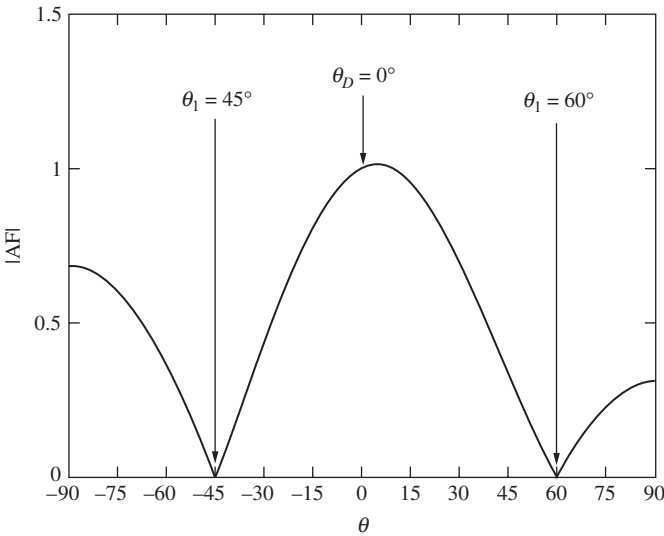


Figure 4.32 Sidelobe cancellation.

There are some limitations to this scheme. The number of nulls cannot exceed the number of array elements. In addition, the array maximum cannot be closer to a null than the array resolution allowed. The array resolution is inversely proportional to the array length.

4.7 Retrodirective Arrays

Retrodirective arrays are the array equivalent of a corner reflector. Corner reflectors can be studied in any standard antenna book. It was Van Atta [21, 22] who invented a scheme to convert a linear array into a reflector. In this case, the array redirects in the incident field in the direction of arrival. Thus, the use of the term “retrodirective” is appropriate. Retrodirective arrays are sometimes also called *self-phasing* arrays, *self-focusing* arrays, *conjugate matched* arrays, or *time-reversal mirrors* [23–25]. A conjugate matched array is retrodirective because it retransmits the signal with the phases conjugated. If the phases are conjugated, it is the same as reversing time in the time domain. This is why retrodirective arrays are sometimes called “time-reversal” arrays. The acoustic community has aggressively pursued time-reversal methods in underwater acoustics as a means for addressing the multipath challenge [26].

It is not necessary that the array is linear in order for the array to be retrodirective. In fact the Van Atta array is a special case of the more general subject of self-phasing arrays. However, in this development, we will restrict our discussion to linear arrays.

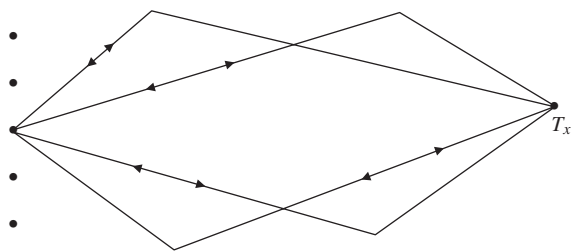


Figure 4.33 Retrodirective array with multipath.

One of the obvious advantages of a retrodirective array is the fact that if the array can redirect energy in the direction of arrival then the array will work extremely well in a multipath environment. Some suggested ways of using a retrodirective array in mobile communications are given by Fusco and Karode [24]. If the same signal arrives from multiple directions, a retrodirective array will retransmit at the same angles and the signal will return to the source as if multipath did not exist. Figure 4.33 shows a linear retrodirective array in a multipath environment.

If indeed the retrodirective array can retransmit along the angles of arrival then the retransmitted signal will retrace the multiple paths back to the transmitter.

4.7.1 Passive retrodirective array

One possible way of implementing a retrodirective array is shown in Fig. 4.34 for an $N = 6$ element array. A plane wave is incident on the array at the angle θ_0 .

The array vector for this N -element array is given as

$$\vec{a} = \left[e^{-j\frac{5}{2}kd \sin \theta} \quad e^{-j\frac{3}{2}kd \sin \theta} \quad \dots \quad e^{j\frac{3}{2}kd \sin \theta} \quad e^{j\frac{5}{2}kd \sin \theta} \right]^T \tag{4.62}$$

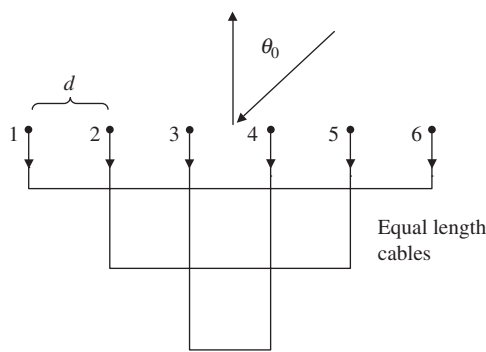


Figure 4.34 Retrodirective array.

The received array vector, at angle θ_0 , is given as

$$\bar{a}_{\text{rec}} = \begin{bmatrix} e^{-j\frac{5}{2}kd \sin \theta_0} & e^{-j\frac{3}{2}kd \sin \theta_0} & \dots & e^{j\frac{3}{2}kd \sin \theta_0} & e^{j\frac{5}{2}kd \sin \theta_0} \end{bmatrix}^T \quad (4.63)$$

The input to element 6, $e^{j\frac{5}{2}kd \sin \theta_0}$, propagates down the transmission line to element 1 and is retransmitted. The same process is repeated for all elements. Thus the transmitted signal for element i was the received signal for element $N - i$. This can be shown to be the same as multiplying the array vector $\bar{a}$ in Eq. (4.62) by the reverse of the vector $\bar{a}_{\text{rec}}$ of Eq. (4.63). Reversing the vector of Eq. (4.63) is the same as reversing or flipping the individual elements. One method of reversing vector elements is through the use of the permutation matrix. In MATLAB, this function can be accomplished through the use of the `fliplr( )` command. The array transmission can be now calculated to be

$$\text{AF} = \begin{bmatrix} e^{-j\frac{5}{2}kd \sin \theta} & e^{-j\frac{3}{2}kd \sin \theta} & \dots & e^{j\frac{3}{2}kd \sin \theta} & e^{j\frac{5}{2}kd \sin \theta} \end{bmatrix}^T \cdot \begin{bmatrix} e^{j\frac{5}{2}kd \sin \theta_0} \\ e^{j\frac{3}{2}kd \sin \theta_0} \\ \vdots \\ e^{-j\frac{3}{2}kd \sin \theta_0} \\ e^{-j\frac{5}{2}kd \sin \theta_0} \end{bmatrix} \quad (4.64)$$

Based on our derivation in Eq. (4.13) this is equivalent to a beamsteered array factor.

$$\text{AF} = \frac{\sin\left(\frac{Nkd}{2}(\sin \theta - \sin \theta_0)\right)}{\sin\left(\frac{kd}{2}(\sin \theta - \sin \theta_0)\right)} \quad (4.65)$$

Thus, this retrodirective array has successfully retransmitted the signal back toward the θ_0 direction. This process works regardless of the angle of arrival (AOA). Therefore, the retrodirective array serves to focus the reflected signal back at the source.

4.7.2 Active retrodirective array

A second method for self-phasing or phase conjugation is achieved through mixing the received signal with a local oscillator. In this case, the local oscillator will be twice the carrier frequency. Although it is not necessary to make this restriction, the analysis is easiest for this case. Each antenna output has its own mixer as shown in Fig. 4.35.

The output of the n th antenna is $\mathcal{V}_n(t, \theta_n)$, given by

$$R_n(t, \theta_n) = \cos(\omega_0 t + \theta_n) \quad (4.66)$$

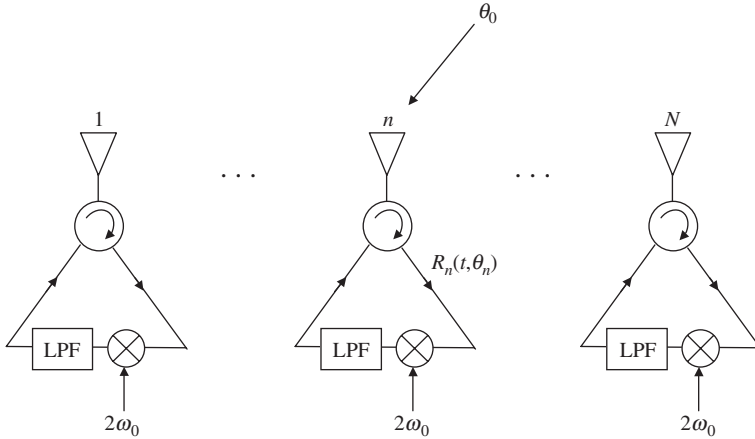


Figure 4.35 Phase conjugation through heterodyne mixing.

The mixer output is given by

$$\begin{aligned} S_{\text{mix}} &= \cos(\omega_0 t + \theta_n) \cdot \cos(2\omega_0 t) \\ &= \frac{1}{2} [\cos(-\omega_0 t + \theta_n) + \cos(3\omega_0 t + \theta_n)] \end{aligned} \quad (4.67)$$

After passing through the low pass filter and selecting the lower sideband, the transmitted signal, for element n , is given by

$$T_n(t, \theta_n) = \cos(\omega_0 t - \theta_n) \quad (4.68)$$

Thus, the phase has been conjugated from the phase imposed at arrival. This array will then redirect the signal back toward the AOA θ_0 . Since, the term in Eq. (4.68) could have been written as $\cos(-\omega_0 t + \theta_n)$. It can thus be seen how this procedure is alternatively called time-reversal and how the array is called a time-reversal mirror. If we choose a different local oscillator frequency, it is sufficient to choose the lower sideband for retransmission.

References

1. Haykin, S., ed., *Array Signal Processing*, Prentice Hall, New York, 1985.
2. Johnson, D., and D. Dudgeon, *Array Signal Processing—Concepts and Techniques*, Prentice Hall, New Jersey, 1993.
3. Trees, H.V., *Optimum Array Processing—Part IV of Detection, Estimation, and Modulation Theory*, Wiley Interscience, New York, 2002.
4. Brookner, E., *Practical Phased-Array Antenna Systems*, Artech House, Boston, MA, 1991.
5. Dudgeon, D.E., "Fundamentals of Digital Array Processing," *Proceedings of IEEE*, Vol. 65, pp. 898–904, June 1977.

6. Balanis, C., *Antenna Theory: Analysis and Design*, 2d ed., Wiley, New York, 1997.
7. Kraus, J.D., and R.J. Marhefka, *Antennas for All Applications*, 3d ed., McGraw-Hill, New York, 2002.
8. Litva, J., and T.K-Y. Lo, *Digital Beamforming in Wireless Communications*, Artech House, 1996.
9. Monzingo, R.A., and T.W. Miller, *Introduction to Adaptive Arrays*, Wiley, New York, 1980.
10. Ertel, R.B., P. Cardieri, K.W. Sowerby, et al., "Overview of Spatial Channel Models for Antenna Array Communication Systems," *IEEE Personal Commun. Mag.*, Vol. 5, No. 1, pp. 10–22, Feb. 1998.
11. Hansen, W.W., and J.R. Woodyard, "A New Principle in Directional Antenna Design," *Proceedings IRE*, Vol. 26, No. 3, pp. 333–345, March 1938.
12. Harris, F.J. "On the Use of Windows for Harmonic Analysis with the DFT," *IEEE Proceedings*, pp. 51–83, Jan. 1978.
13. Nuttall, A.H., in "Some Windows with Very Good Sidelobe Behavior," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. ASSP-29, No. 1, Feb. 1981.
14. Johnson, R.C., and H. Jasik, *Antenna Engineering Handbook*, 2d ed., McGraw-Hill, New York, pp. 20–16, 1984.
15. Mailloux, R.J., *Phased Array Antenna Handbook*, Artech House, Norwood, MA, 1994.
16. Hansen, R.C., *Phased Array Antennas*, Wiley, New York, 1998.
17. Pattan, B., *Robust Modulation Methods and Smart Antennas in Wireless Communications*, Prentice Hall, New York, 2000.
18. Butler, J., and R. Lowe, "Beam-Forming Matrix Simplifies Design of Electrically Scanned Antennas," *Electronic Design*, April 12, 1961.
19. Shelton, J.P., and K.S. Kelleher, "Multiple Beams From Linear Arrays," *IRE Transactions on Antennas and Propagation*, March 1961.
20. Howells, P.W., "Intermediate Frequency Sidelobe Canceller," U.S. Patent 3202990, Aug. 24, 1965.
21. Van Atta, L.C., "Electromagnetic Reflector," U.S. Patent 2908002, Oct. 6, 1959.
22. Sharp, E.D., and M.A. Diab, "Van Atta Reflector Array," *IRE Transactions on Antennas and Propagation Communications*, Vol. AP-8, pp. 436–438; July 1960.
23. Skolnik, M.I. and D.D. King, "Self-Phasing Array Antennas," *IEEE Transactions on Antennas and Propagation*, Vol. AP-12, No. 2, pp. 142–149, March 1964.
24. Fusco, V.F., and S.L. Karode, "Self-Phasing Antenna Array Techniques for Mobile Communications Applications," *Electronics and Communication Engineering Journal*, Dec. 1999.
25. Blomgren, P., G. Papanicolaou, and H. Zhao, "Super-Resolution in Time-Reversal Acoustics," *Journal Acoustical Society of America*, Vol. 111, No. 1, Pt. 1, Jan. 2002.
26. Fink, M., "Time-Reversed Acoustics," *Scientific American*, pp. 91–97, Nov. 1999.

Problems

4.1 Use Eq. (4.14) and MATLAB and plot the array factor in rectangular coordinate form for the following broadside array cases

- (a) $N = 4, d = \lambda/2$
- (b) $N = 8, d = \lambda/2$
- (c) $N = 8, d = \lambda$

4.2 For the three arrays given in Prob. 4.1, calculate θ_{null} and θ_s for each of the arrays given above.

4.3 Repeat Prob. 4.1 above but for end-fire arrays.

4.4 Use MATLAB and the command `trapz()` to calculate the maximum directivity for the following two arrays:

- (a) Broadside with $N = 8$, $d = \lambda/2$
- (b) End-fire with $N = 8$, $d = \lambda/2$

4.5 Use MATLAB and the command `trapz()` to calculate the maximum directivity for the beamsteered array where $d = \lambda/2$, $N = 8$.

- (a) $\theta_0 = 30^\circ$
- (b) $\theta_0 = 45^\circ$

4.6 What is the beamwidth for the following array parameters?

- (a) $\theta_0 = 0^\circ$, $N = 8$, $d = \lambda/2$
- (b) $\theta_0 = 45^\circ$, $N = 8$, $d = \lambda/2$
- (c) $\theta_0 = 90^\circ$, $N = 8$, $d = \lambda/2$

4.7 For an $N = 6$, $d = \lambda/2$ uniformly weighted broadside array, plot the array factor. Superimpose plots of the same array with the following weights. (Create one new plot for each new set of weights)

- (a) Kaiser-Bessel for $\alpha = 2$ using `kaiser(N,  $\alpha$ )`
- (b) Blackman-Harris using `blackmanharris(N)`
- (c) Nuttall using `nuttallwin(N)`
- (d) Chebyshev window for $R = 30$ dB using `chebwin(N, R)`

4.8 Repeat Prob. 4.7 for $N = 9$, $d = \lambda/2$.

4.9 Using MATLAB, create and superimpose three normalized array factor plots using the `chebwin()` function for $R = 20$, 40, and 60 dB. $N = 9$, $d = \lambda/2$

4.10 Using MATLAB, create and superimpose three normalized array factor plots using the `chebwin()` function for $R = 40$ and beamsteer the array to three angles such that $\theta_0 = 0^\circ$, 30° , 60° . $N = 9$, $d = \lambda/2$.

4.11 For $d = \lambda/2$, use MATLAB and plot the broadside array beamwidth vs. element number N for $2 < N < 20$.

4.12 For $d = \lambda/2$, use MATLAB and plot the $N = 8$ element array beamwidth for a range of steering angles such that $0 < \theta_0 < 90^\circ$.

4.13 For the $N = 40$ element circular array with radius $a = 2\lambda$. Use MATLAB to plot the elevation plane pattern in the x - z plane when the array is beamsteered to $\theta_0 = 20^\circ$, 40° , and 60° , $\phi_0 = 0$. Plot for $(0 < \theta < 180^\circ)$

4.14 For the $N = 40$ element circular array with radius $a = 2\lambda$. Use MATLAB to plot the azimuth plane pattern in the x - y plane when the array is beamsteered to $\theta_0 = 90^\circ$, $\phi_0 = 0^\circ$, 40° , and 60° . Plot for $(-90^\circ < \phi < 90^\circ)$

4.15 Design and plot a 5×5 element array with equal element spacing such that $d_x = d_y = .5\lambda$. Let the array be beamsteered to $\theta_0 = 45^\circ$ and $\phi_0 = 90^\circ$.

The element weights are chosen to be the Blackman-Harris weights using the `blackmanharris( )` command in MATLAB. Plot the pattern for the range $0 \leq \theta \leq \pi/2$ and $0 \leq \phi \leq 2\pi$.

4.16 Use Eq. (4.56) to create scalloped beams for the $N = 6$ -element array with $d = \lambda/2$.

- (a) What are the ν values?
- (b) What are the angles of the scalloped beams?
- (c) Plot and superimpose all beams on a polar plot similar to Fig. 4.29

4.17 For the fixed beam sidelobe canceller in Sec. 4.8, with $N = 3$ -antenna elements, calculate the array weights to receive the desired signal at $\theta_D = 30^\circ$, and to suppress the interfering signals arriving at $\theta_1 = -30^\circ$ and $\theta_2 = -60^\circ$.

Principles of Random Variables and Processes

Every wireless communication system or radar system must take into account the noise-like nature of the arriving signals as well as the internal system noise. The arriving intelligent signals are usually altered by propagation, spherical spreading, absorption, diffraction, scattering, and/or reflection from various objects. This being the case, it is important to know the statistical properties of the propagation channel as well as the statistical properties of the noise internal to the system. Chapter 6 will address the subject of multipath propagation, which will be seen to be a random process. Chapters 7 and 8 will deal with system noise as well as the statistical properties of arriving signals. In addition, methodologies used in the remaining chapters will require computations based upon the assumption that the signals and noise are random.

It is thus assumed that students or researchers using this book have a working familiarity with random processes. Typically this material is covered in any undergraduate course in statistical topics or communications. However, for the purposes of consistency in this text, we will perform a brief review of some fundamentals of random processes. Subsequent chapters will apply these principles to specific problems.

Several books are devoted to the subject of random processes including texts by Papoulis [1], Peebles [2], Thomas [3], Schwartz [4], and Haykin [5]. The treatment of the subject in this chapter is based upon the principles discussed in depth in these references.

5.1 Definition of Random Variables

In the context of communication systems, the received voltages, currents, phases, time delays, and angles-of-arrival tend to be random

variables. As an example, if one were to conduct measurements of the receiver phase every time the receiver is turned on, the numbers measured would tend to be randomly distributed between 0 and 2π . One could not say with certainty what value the next measurement will produce but one could state the probability of getting a certain measured value. A random variable is a function that describes all possible outcomes of an experiment. In general, some values of a random variable are more likely to be measured than other values. The probability of getting a specific number when rolling a die is equal to the probability of any other number. However, most random variables in communications problems do not have equally likely probabilities. Random variables can either be discrete or continuous variables. A random variable is discrete if the variable can only take on a finite number of values during an observation interval. An example of a discrete random variable might be the arrival angle for indoor multipath propagation. A random variable is continuous if the variable can take on a continuum of values during an observation interval. An example of a continuous random variable might be the voltage associated with receiver noise or the phase of an arriving signal. Since random variables are the result of random phenomena, it is often best to describe the behavior of random variables using probability density functions.

5.2 Probability Density Functions

Every random variable x is characterized by a probability density function $p(x)$. The *probability density function* (pdf) is established after a large number of measurements have been performed, which determine the likelihood of all possible values of x . A discrete random variable possesses a discrete pdf. A continuous random variable possesses a continuous pdf. Figure 5.1 shows a typical pdf for a discrete random variable. Figure 5.2 shows a typical pdf for a continuous random variable.

The probability that x will take on a range of values between two limits x_1 and x_2 is defined by

$$P(x_1 \leq x \leq x_2) = \int_{x_1}^{x_2} p(x) dx \quad (5.1)$$

There are two important properties for pdfs. First, no event can have a negative probability. Thus

$$p(x) \geq 0 \quad (5.2)$$

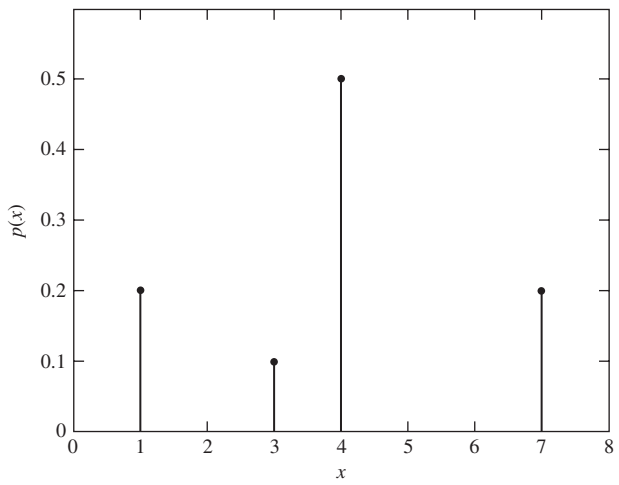


Figure 5.1 Probability density function for discrete x values.

Second, the probability that an x value exists somewhere over its range of values is certain. Thus

$$\int_{-\infty}^{\infty} p(x) dx = 1 \tag{5.3}$$

Both properties must be satisfied by any pdf. Since the total area under the pdf is equal to 1, the probability of x existing over a finite range of possible values is always less than 1.

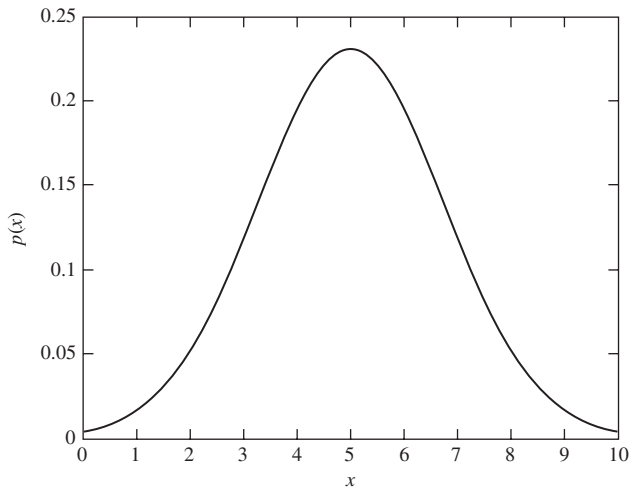


Figure 5.2 Probability density function for continuous x values.

5.3 Expectation and Moments

It is valuable to understand various properties of the random variable x or various properties of functions of the random variable x . The most obvious property is the statistical average. The *statistical average* is defined as the expected value denoted by E . Thus, the expected value of x is defined as

$$E[x] = \int_{-\infty}^{\infty} xp(x) dx \quad (5.4)$$

Not only can we find the expected value of x but we can also find the expected value of any function of x . Thus

$$E[f(x)] = \int_{-\infty}^{\infty} f(x)p(x) dx \quad (5.5)$$

The function of x could be x^2 , x^3 , $\cos(x)$, or any other operation on the random variable. The expected value of x is typically called the *first moment* denoted as m_1 .

$$m_1 = \int_{-\infty}^{\infty} xp(x) dx \quad (5.6)$$

The n th *moment* is defined as the expected value of x^n , thus,

$$m_n = \int_{-\infty}^{\infty} x^n p(x) dx \quad (5.7)$$

The concept of moments is borrowed from the terminology of moments in mechanics.

If the random variable was expressed in volts, the first moment would correspond to the average, mean, or dc voltage. The second moment would correspond to the average power.

The spreading about the first moment is called the *variance* and is defined as

$$E[(x - m_1)^2] = \mu_2 = \int_{-\infty}^{\infty} (x - m_1)^2 p(x) dx \quad (5.8)$$

The *standard deviation* is denoted by σ and is defined as the spread about the mean, thus,

$$\sigma = \sqrt{\mu_2} \quad (5.9)$$

By expanding the squared term in Eq. (5.8) it can be shown that $\sigma = m_2 - m_1^2$.

The first moment and standard deviation tend to be the most useful descriptors of the behavior of a random variable. However, other

moments may need to be calculated to understand more fully the behavior of a given random variable x . Since the calculation of each new moment requires a reevaluation Eq. (5.7), sometimes it is useful to utilize the moment generating function that will simplify the calculation of multiple moments. The *moment generating function* is defined as

$$E[e^{sx}] = \int_{-\infty}^{\infty} e^{sx} p(x) dx = F(s) \quad (5.10)$$

The moment generating function resembles the Laplace transform of the pdf. This now brings us to the moment theorem. If we differentiate Eq. (5.10) n times with respect to s , it can be shown that

$$F^n(s) = E[x^n e^{sx}] \quad (5.11)$$

Thus, when $s = 0$, we can derive the n th moment as shown in Eq. (5.12)

$$F^n(0) = E[x^n] = m_n \quad (5.12)$$

Example 5.1 If the discrete pdf of the random variable is given by $p(x) = .5[\delta(x + 1) + \delta(x - 1)]$, what are the first three moments using the moment generating function $F(s)$?

Solution Finding the moment generating function we have

$$\begin{aligned} F(s) &= \int_{-\infty}^{\infty} e^{sx} (.5)[\delta(x + 1) + \delta(x - 1)] dx = .5[e^{-s} + e^s] \\ &= \cosh(s) \end{aligned}$$

The first moment is given as

$$m_1 = F^1(s)|_{s=0} = \left. \frac{d \cosh(s)}{ds} \right|_{s=0} = \sinh(0) = 0$$

The second moment is given as

$$m_2 = F^2(s)|_{s=0} = \cosh(0) = 1$$

The third moment is given as

$$m_3 = F^3(s)|_{s=0} = \sinh(0) = 0$$

5.4 Common Probability Density Functions

There are numerous pdfs commonly used in both radar, sonar, and communications. These pdfs describe the characteristics of the receiver noise, the arriving signal from multipaths, the distribution of the phase, envelope, and power of arriving signals. A quick summary of these pdfs

and their behavior will be useful in supporting concepts addressed in Chaps. 6, 7, and 8.

5.4.1 Gaussian density

The Gaussian or normal probability density is perhaps the most common pdf. The Gaussian distribution generally defines the behavior of noise in receivers and also the nature of the random amplitudes of arriving multipath signals. According to the Central Limit Theorem, the sum of numerous continuous random variables as the number increases, tends toward a Gaussian distribution. The *Gaussian density* is defined as

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(x-x_0)^2}{2\sigma^2}} \quad -\infty \leq x \leq \infty \quad (5.13)$$

This is a bell-shaped curve, which is symmetric about the mean value x_0 , and has a standard deviation of σ . A plot of a typical Gaussian distribution is shown in Fig. 5.3.

Example 5.2 For the Gaussian probability density function with $x_0 = 0$ and $\sigma = 2$, calculate the probability that x will exist over the range $0 \leq x \leq 4$.

Solution Invoking Eqs. (5.1) and (5.13) we can find the probability

$$P(0 \leq x \leq 4) = \int_0^4 \frac{1}{\sqrt{8\pi}} e^{-x^2/8} = .477$$

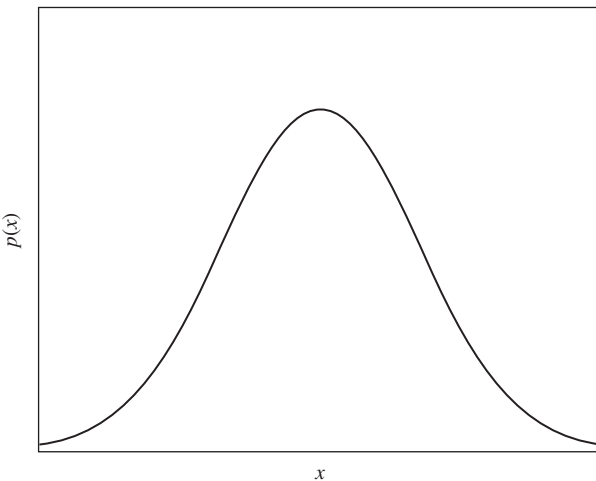


Figure 5.3 Gaussian density function.

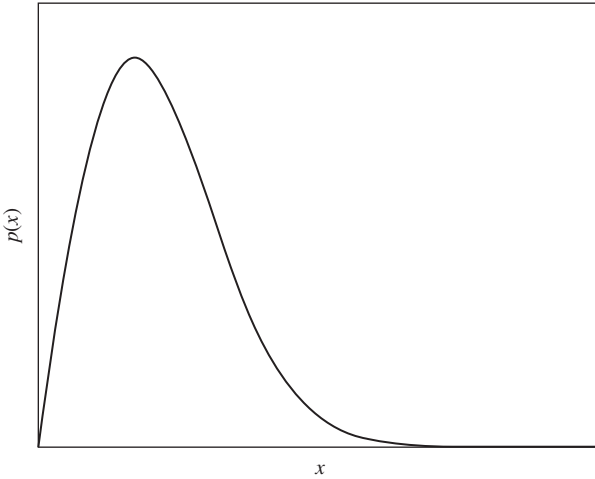


Figure 5.4 Rayleigh density function.

5.4.2 Rayleigh density

The Rayleigh probability density generally results when one finds the envelope of two independent Gaussian processes. This envelope can be found at the output of a linear filter where the inputs are Gaussian random variables. Rayleigh distributions are normally attributed to the envelope of multipath signals when there is no direct path. The *Rayleigh distribution* is defined as

$$p(x) = \frac{x}{\sigma^2} e^{-x^2/2\sigma^2} \quad x \geq 0 \quad (5.14)$$

The standard deviation can be shown to be σ . A plot of a typical Rayleigh distribution is shown in Fig. 5.4.

Example 5.3 For the Rayleigh probability density function with $\sigma = 2$, calculate the probability that x will exist over the range $0 \leq x \leq 4$.

Solution Invoking Eqs. (5.1) and (5.14) we can find the probability

$$P(0 \leq x \leq 4) = \int_0^4 \frac{x}{4} e^{-x^2/8} = .865$$

5.4.3 Uniform density

The uniform distribution is normally attributed to the distribution of the random phase for propagating signals. Not only does the phase delay tend to be uniformly distributed but often the angles of arrival

for diverse propagating waves can also take on a uniform distribution. The *uniform distribution* is defined as

$$p(x) = \frac{1}{b-a} [u(x-a) - u(x-b)] \quad a \leq x \leq b \quad (5.15)$$

This mean value can be shown to be $(a+b)/2$. A plot of a typical uniform distribution is shown in Fig. 5.5.

Example 5.4 For the uniform distribution where $a = -2$, and $b = 2$, use the moment generating function to find the first three moments.

Solution By substituting Eq. (5.15) into Eq. (5.10) we have

$$\begin{aligned} F(s) &= \frac{1}{4} \int_{-\infty}^{\infty} [u(x+2) - u(x-2)] e^{sx} dx = \frac{1}{4} \int_{-2}^2 e^{sx} dx \\ &= \frac{1}{2s} \sinh(2s) \end{aligned}$$

The first moment is found as

$$m_1 = F'(s)|_{s=0} = \frac{\cosh(2s)}{s} - \frac{\sinh(2s)}{2s^2} = 0$$

The second moment is found as

$$m_2 = F''(s)|_{s=0} = \sinh(2s) \left[\frac{2}{s} + \frac{1}{s^3} \right] - \frac{2 \cosh(2s)}{s^2} = \frac{4}{3}$$

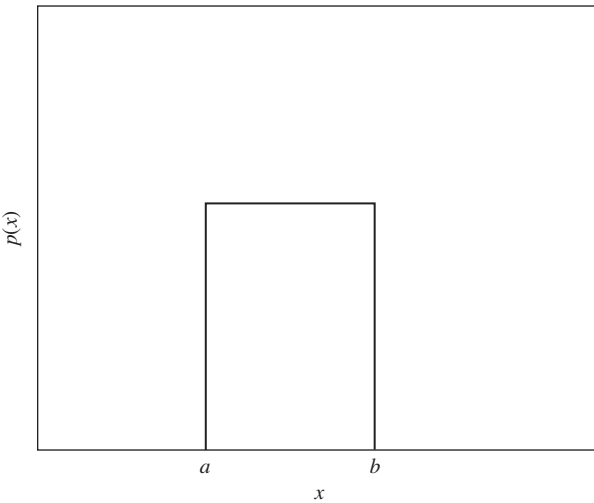


Figure 5.5 Uniform density function.

The third moment is found as

$$m_3 = F^3(s)|_{s=0} = \cosh(2s) \left[\frac{4}{s} + \frac{6}{s^3} \right] - \sinh(2s) \left[\frac{6}{s^2} + \frac{3}{s^4} \right] = 0$$

5.4.4 Exponential density

The exponential density function is sometimes used to describe the angles of arrival for incoming signals. It can also be used to describe the power distribution for a Rayleigh process. The *Exponential density* is the Erlang density when $n = 1$ ([1]) and is defined by

$$p(x) = \frac{1}{\sigma} e^{-x/\sigma} \quad x \geq 0 \quad (5.16)$$

The mean value can be shown to be σ . The standard deviation can also be shown to be σ . The literature sometimes replaces σ with $2\sigma^2$ in Eq. (5.16). A plot of a typical exponential distribution is shown in Fig. 5.6.

Example 5.5 Using the exponential density with $\sigma = 2$, calculate the probability that $2 \leq x \leq 4$ and find the first two moments using the moment generating function.

Solution The probability is given as

$$P(2 \leq x \leq 4) = \int_2^4 \frac{1}{\sigma} e^{-x(1/\sigma)} dx = .233$$

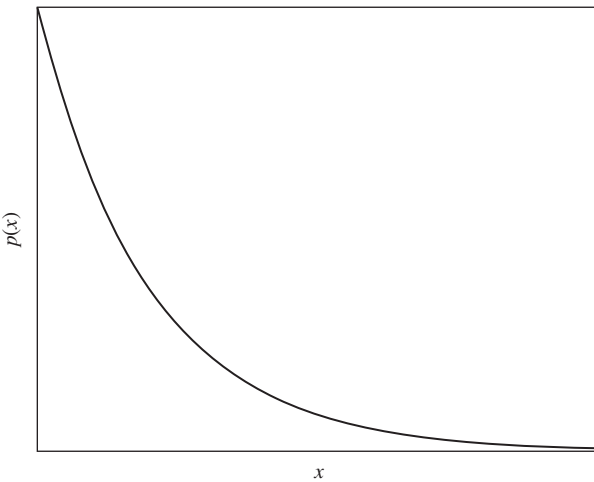


Figure 5.6 Exponential density function.

The moment generating function is derived as

$$F(s) = \int_0^{\infty} \frac{1}{\sigma} e^{-x(\frac{1}{\sigma} - s)} dx = \frac{1}{1 - \sigma s}$$

Finding the first moment

$$F^1(s) = \frac{\sigma}{(1 - \sigma s)^2} \Big|_{s=0} = \sigma = 2$$

Finding the second moment

$$F^{21}(s) = \frac{2\sigma^2}{(1 - \sigma s)^3} \Big|_{s=0} = 2\sigma^2 = 8$$

5.4.5 Rician density

The Rician distribution is common for propagation channels where there is a direct path signal added to the multipath signals. The direct path inserts a nonrandom carrier thereby modifying the Rayleigh distribution. Details of the derivation of the Rician distribution can be found in [1, 4]. The *Rician distribution* is defined as

$$p(x) = \frac{x}{\sigma^2} e^{-\frac{(x^2 + A^2)}{2\sigma^2}} I_0 \left(\frac{xA}{\sigma^2} \right) \quad x \geq 0, A \geq 0 \quad (5.17)$$

where $I_0(\cdot)$ is the Modified Bessel function of first kind and zero-order.

A plot of a typical Rician distribution is shown in Fig. 5.7.

Example 5.6 For the Rician distribution with $\sigma = 2$ and $A = 2$, what is the

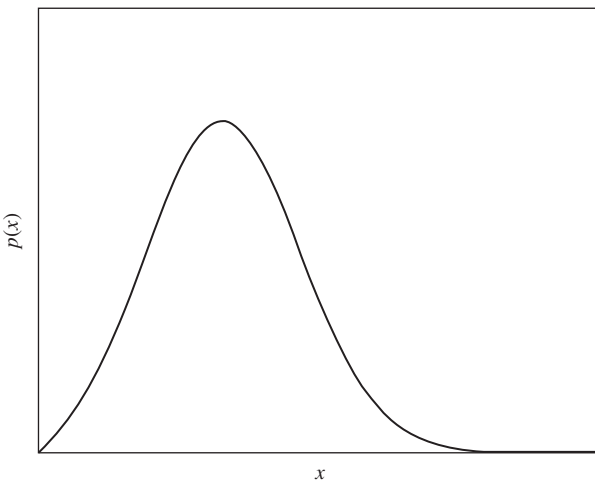


Figure 5.7 Rician density function.

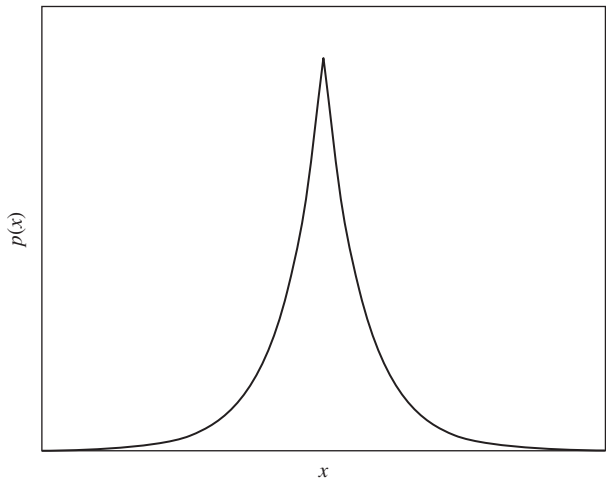


Figure 5.8 Laplace density function.

probability that $x \geq 5$?

Solution Using Eqs. (5.17) and (5.1) we have the probability as

$$P(x \geq 5) = \int_5^\infty \frac{x}{4} e^{-\frac{(x^2+4)}{8}} I_0\left(\frac{x}{2}\right) dx = .121$$

5.4.6 Laplace density

The Laplace density function is generally attributed to the distribution of indoor or congested urban angles of arrival. The Laplace distribution is given as

$$p(x) = \frac{1}{\sqrt{2}\sigma} e^{-|\frac{\sqrt{2}x}{\sigma}|} \quad -\infty \leq x \leq \infty \tag{5.18}$$

Since the Laplace distribution is symmetric about the origin, the first moment is zero. The second moment can be shown to be σ^2 . A plot of the Laplace distribution is shown in Fig. 5.8.

There are many other probability density functions described in the literature but the six functions mentioned in this chapter are the most common distribution functions applied to wireless communication problems.

5.5 Stationarity and Ergodicity

In realistic applications, we may know the statistical properties of the signals and noise but we are often confronted with the challenge of performing operations on a limited block of sampled data. If the statistical mean m_1 is the average value of the random variable x , one might

intuitively assume that the time average would be equal to the statistical average. We can estimate the statistical average by using the time average over a block length T . The time average for the random variable x can be written as

$$\hat{x} = \frac{1}{T} \int_0^T x(t) dt \quad (5.19)$$

where $\hat{x}$ is the estimate of the statistical average of x .

If the data is sampled data, Eq. (5.19) can be rewritten as a series to be expressed as

$$\hat{x} = \frac{1}{K} \sum_{k=1}^K x(k) \quad (5.20)$$

Since the random variable $x(t)$ is changing with time, one might expect that the estimate $\hat{x}$ might also vary with time depending on the block length T . Since we are performing a linear operation on the random variable x , we will produce a new random variable $\hat{x}$. One might expect that the time average and the statistical average would be similar if not identical. We can take the expected value of both sides of Eq. (5.19)

$$E[\hat{x}] = \frac{1}{T} \int_0^T E[x(t)] dt \quad (5.21)$$

If all statistics of the random variable x do not change with time, the random process is said to be *strict-sense stationary* [1]. A strict sense stationary process is one in which the statistical properties are invariant to a shift in the time origin. If the mean value of a random variable does not change with time, the process is said to be *wide-sense stationary*. If x is wide-sense stationary, Eq. (5.21) simplifies to

$$E[\hat{x}] = E[x(t)] = m_1 \quad (5.22)$$

In reality, the statistics might change for short blocks of time T but stabilize over longer blocks of time. If by increasing T (or K) we can force the time average estimate to converge to the statistical average, the process is said to be *ergodic* in the mean or *mean-ergodic*. This can be written as

$$\lim_{T \rightarrow \infty} \hat{x} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x(t) dt = m_1 \quad (5.23)$$

or

$$\lim_{T \rightarrow \infty} \hat{x} = \lim_{T \rightarrow \infty} \frac{1}{K} \sum_{k=1}^K x(k) = m_1 \quad (5.24)$$

In a similar way, we can also use a time average to estimate the *variance* of x defined as

$$\hat{\sigma}_x^2 = \frac{1}{T} \int_0^T (x(t) - \hat{x})^2 dt \quad (5.25)$$

If the data is sampled data, Eq. (5.25) can be rewritten as a series to be expressed as

$$\hat{\sigma}_x^2 = \frac{1}{K} \sum_{k=1}^K (x(k) - \hat{x})^2 \quad (5.26)$$

If by increasing T (or K) we can force the variance estimate to converge to the statistical variance, the process is said to be *ergodic* in the variance or *variance-ergodic*. This can be written as

$$\lim_{T \rightarrow \infty} \hat{\sigma}_x^2 = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T (x(t) - \hat{x})^2 dt = \sigma_x^2 \quad (5.27)$$

or

$$\lim_{T \rightarrow \infty} \hat{\sigma}_x^2 = \lim_{T \rightarrow \infty} \frac{1}{K} \sum_{k=1}^K (x(k) - \hat{x})^2 = \sigma_x^2 \quad (5.28)$$

In summary, stationary processes are ones in which the statistics of the random variables do not change at different times. Ergodic processes are ones where it is possible to estimate the statistics, such as mean, variance, and autocorrelation, from the measured values in time. Stationarity and ergodicity will prove valuable in practical communication systems because, under certain conditions, one can reliably estimate the mean, variance, and other parameters based upon computing the time averages.

5.6 Autocorrelation and Power Spectral Density

It is valuable to know how well a random variable correlates with itself at different points in time. That is, how does x at the time t_1 correlate with x at the time t_2 ? This correlation is defined as an *autocorrelation* since we are correlating x with itself. The autocorrelation is normally written as

$$R_x(t_1, t_2) = E[x(t_1)x(t_2)] \quad (5.29)$$

If the random variable x is wide-sense stationary, the specific values of t_1 and t_2 are not as important as the time interval between these

two values defined by τ . Thus, for a wide-sense stationary process, the autocorrelation can be rewritten as

$$R_x(\tau) = E[x(t)x(t + \tau)] \quad (5.30)$$

It should be noted that the autocorrelation value at $\tau = 0$ is the second moment. Thus,

$$R_x(0) = E[x^2] = m_2.$$

Again, in practical systems where we are constrained to process limited blocks of data, one is forced to estimate the autocorrelation based upon using a time average. Therefore, the *estimate* of the autocorrelation can be defined as

$$\hat{R}_x(\tau) = \frac{1}{T} \int_0^T x(t)x(t + \tau) dt \quad (5.31)$$

If the data is sampled data, Eq. (5.31) can be rewritten as a series to be expressed as

$$\hat{R}_x(n) = \frac{1}{K} \sum_{k=1}^K x(k)x(k + n) \quad (5.32)$$

If by increasing T (or K) we can force the autocorrelation estimate to converge to the statistical autocorrelation, the process is said to be *ergodic* in the autocorrelation or *autocorrelation-ergodic*. This can be written as

$$\lim_{T \rightarrow \infty} \hat{R}_x(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x(t)x(t + \tau) dt = R_x(\tau) \quad (5.33)$$

It should be noted that the units of the autocorrelation function for electrical systems are normally expressed in watts. Thus $R_x(0)$ yields the average power of the random variable x .

As with normal signals and linear systems, it is instructive to understand the behavior of the spectrum of the random variable x . Such parameters as bandwidth and center frequency help the system designer to understand how to best process the desired signal. The autocorrelation itself is a function of the time delay between two time-separated random variables. Thus, the autocorrelation is subject to Fourier analysis. Let us define the power spectral density as the Fourier transform of the autocorrelation function.

$$S_x(f) = \int_{-\infty}^{\infty} R_x(\tau)e^{-j2\pi f\tau} d\tau \quad (5.34)$$

$$R_x(\tau) = \int_{-\infty}^{\infty} S_x(f)e^{j2\pi f\tau} df \quad (5.35)$$

The Fourier transform pair in Eqs. (5.34) and (5.35) is frequently referred to as the *Wiener-Khinchin pair* [6].

5.7 Correlation Matrix

In the previous treatment of random variables, we assumed that only one random variable x existed and we performed expectation operations on these scalar values. Several circumstances arise when a collection of random variables exist. One such example is the output of each element of an antenna array. If an incoming planewave induces a random voltage on all M -array elements, the received signal x is a vector. Using the notation of Chap. 4 we can describe the array element output voltages for one incident planewave.

$$\tilde{x}(t) = \tilde{a}(\theta) \cdot s(t) \quad (5.36)$$

where $s(t)$ = incident monochromatic signal at time t
 $\tilde{a}(\theta)$ = M -element array steering vector for the θ direction of arrival.

Let us now define the $M \times M$ -array correlation matrix $\bar{R}_{xx}$ as

$$\begin{aligned} \bar{R}_{xx} &= E[\tilde{x} \cdot \tilde{x}^H] = E[(\tilde{a}s)(s^* \tilde{a}^H)] \\ &= \tilde{a} E[|s|^2] \tilde{a}^H \\ &= S \tilde{a} \cdot \tilde{a}^H \end{aligned} \quad (5.37)$$

where $()^H$ indicates the Hermitian transpose and $S = E[|s|^2]$.

The correlation matrix in Eq. (5.37) assumes that we are calculating the ensemble average using the expectation operator $E[]$. It should be noted that this is not a vector autocorrelation because we have imposed no time delay in the vector $\tilde{x}$.

For realistic systems where we have a finite data block, we must resort to estimating the correlation matrix using a time average. Therefore, we can re-express the operation in Eq. (5.37) as

$$\hat{R}_{xx} = \frac{1}{T} \int_0^T \tilde{x}(t) \cdot \tilde{x}(t)^H dt = \frac{\tilde{a} \cdot \tilde{a}^H}{T} \int_0^T |s(t)|^2 dt \quad (5.38)$$

If the data is sampled data, Eq. (5.38) can be rewritten as a series to be expressed as

$$\hat{R}_{xx} = \frac{\tilde{a} \cdot \tilde{a}^H}{K} \sum_{k=1}^K |s(k)|^2 \quad (5.39)$$

If by increasing T (or K) we can force the correlation matrix estimate to converge to the statistical correlation matrix, the process is said to be *ergodic* in the correlation matrix. This can be written as

$$\lim_{T \rightarrow \infty} \hat{R}_{xx}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x(t)x(t)^H dt = \bar{R}_{xx} \quad (5.40)$$

The correlation matrix will be heavily used in Chaps. 7 and 8.

References

1. Papoulis, A., *Probability, Random Variables, and Stochastic Processes*, 2nd ed., McGraw-Hill, New York, 1984.
2. Peebles, P., *Probability, Random Variables, and Random Signal Principles*, McGraw-Hill, New York, 1980.
3. Thomas, J., *An Introduction to Statistical Communication Theory*, Wiley, New York, 1969.
4. Schwartz, M., *Information, Transmission, Modulation, and Noise*, McGraw-Hill, New York, 1970.
5. Haykin, S., *Communication Systems*, 2nd ed., Wiley, New York, 1983.
6. Papoulis, A., and S. Pillai, *Probability, Random Variables, and Stochastic Processes*, 4th ed., McGraw-Hill, New York, 2002.

Problems

5.1 For the discrete probability density function given as

$$p(x) = \frac{1}{3}[\delta(x) + \delta(x-1) + \delta(x-2)]$$

- (a) What is the moment generating function $F(s)$
- (b) Calculate the first two moments using Eq. (5.7)
- (c) Calculate the first two moments using the moment generating function and Eq. (5.12)

5.2 For the Gaussian density with $\sigma = 1$, and $x_0 = 3$

- (a) Use MATLAB and plot the function for $-10 \leq x \leq 10$.
- (b) What is the probability $P(x \leq 2)$?

5.3 For the Rayleigh density with $\sigma = 2$

- (a) Use MATLAB and plot the function for $0 \leq x \leq 10$.
- (b) What is the probability $P(x \geq 2)$?

5.4 For the uniform density with $a = 0$, $b = 5$

- (a) Find the first two moments using Eq. (5.7)
- (b) Find the first two moments using Eq. (5.12)

5.5 For the exponential density with $\sigma = 2$

- (a) Use MATLAB and plot the function for $0 \leq x \leq 5$.
- (b) What is the probability $P(x \geq 2)$?

- 5.6** For the Rician density with $\sigma = 3$ and $A = 5$
- Use MATLAB and plot the function for $0 \leq x \leq 10$.
 - What is the probability $P(x \geq 5)$?
- 5.7** For the Laplace density with $\sigma = 2$
- Use MATLAB and plot the function for $-6 \leq x \leq 6$.
 - What is the probability $P(x \leq -2)$?
- 5.8** Create the 30 sample zero mean Gaussian random variable x using MATLAB such that the $\sigma = 2$. This is done by the command $x = \sigma * \text{randn}(1, 30)$. This is a discrete time series of block length $K = 30$.
- Use Eq. (5.20) to estimate the mean value
 - Use Eq. (5.25) to estimate the standard deviation σ_x
 - What is the percent error between these estimates and the mean and standard deviation for the true Gaussian process?
- 5.9** Use the same sequence from Prob. 8
- Use the MATLAB `xcorr()` command to calculate and plot the autocorrelation $R_x(n)$ of x .
 - Use the FFT command and `fftshift` to calculate the power spectral density of $R_x(n)$. Plot the absolute value.
- 5.10** For the $N = 2$ element array with elements spaced $\lambda/2$ apart
- What is the array steering vector for the $\theta = 30^\circ$?
 - Define the time signal impinging upon the array as $s(t) = 2 \exp(\pi t/T)$. Use Eq. (5.38) and calculate the array correlation matrix.

This page intentionally left blank

Propagation Channel Characteristics

Free-space transmission occurs when the received signal is exclusively the result of direct path propagation. In this case there is no interference at the receiver caused by multipath signals. The received signal strength calculations are straightforward and deterministic. The free-space transmission model is a useful construct which can be used to understand fundamental propagation behavior. However, the free-space model is unrealistic because it fails to account for the numerous terrestrial effects of multipath propagation. It is normally assumed in this chapter that the propagation channel includes at least two propagation paths.

A *channel* is defined as the communication path between transmit and receive antennas. The channel accounts for all possible propagation paths as well as the effects of absorption, spherical spreading, attenuation, reflection losses, Faraday rotation, scintillation, polarization dependence, delay spread, angular spread, Doppler spread, dispersion, interference, motion, and fading. It may not be necessary that any one channel has all of the above effects but often channels have multiple influences on communication waveforms. Obviously, the complexity of the channel increases as the number of available propagation paths increases. It also becomes more complex if one or more variables vary with time such as the receiver or transmitter position. Several excellent texts exist which describe channel characteristics [1–9].

Indoor propagation modeling can be a formidable challenge. This is partly due to the regular and periodic location of structures such as windows, doors, wall studs, ceiling tiles, electrical conduits, ducts, and plumbing. This is also partly due to the very close proximity of

scattering objects relative to the transmitter and/or receiver. An excellent treatment of indoor propagation channels can be found in Sarkar et al. [5], Shankar [3], and Rappaport [9]. This chapter will preview channel characterization basics but will assume outdoor propagation conditions.

6.1 Flat Earth Model

In Chap. 2 we discussed propagation over flat earth. This propagation model was simplistic but it demonstrated the elements of more complex propagation scenarios. We will repeat Fig. 2.17 as Fig. 6.1. The transmitter is at height h_1 while the receiver is at height h_2 . The direct path term has path length r_1 . There is also a reflected term due to the presence of a ground plane, this is called the *indirect path*. Its overall path length is r_2 . The ground causes a reflection at the point y with a reflection coefficient R . The reflection coefficient R is normally complex and can alternatively be expressed as $R = |R|e^{j\psi}$. The reflection coefficient can be found from the Fresnel reflection coefficients as described in Eq. (2.63) or (2.69). The reflection coefficient is polarization dependent. One expression is used for parallel polarization where the E field is parallel to the plane of incidence (E is perpendicular to the ground). The other expression is for perpendicular polarization where the E field is perpendicular to the plane of incidence (E is parallel to the ground).

Through simple algebra, it can be shown that

$$r_1 = \sqrt{d^2 + (h_2 - h_1)^2} \tag{6.1}$$

$$r_2 = \sqrt{d^2 + (h_2 + h_1)^2} \tag{6.2}$$

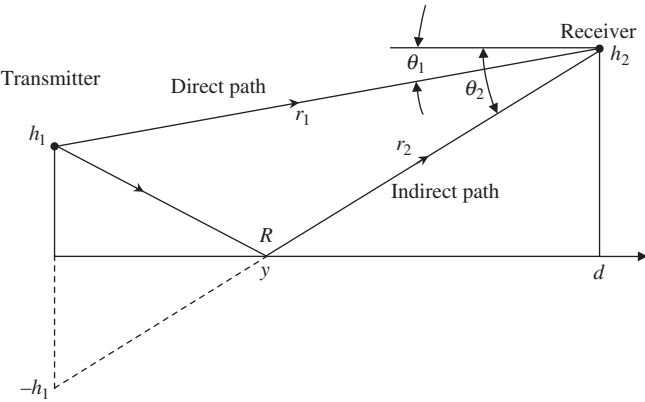


Figure 6.1 Propagation over flat earth.

The total received phasor field can be found to be

$$E_{rs} = \frac{E_0 e^{-jk r_1}}{r_1} + \frac{E_0 \text{Re} e^{-jk r_2}}{r_2} \quad (6.3)$$

where E_{rs} = phasor representation of the received electric field E_r .

The reflection point is given by $y = dh_1/(h_1 + h_2)$. Therefore, the angles of arrival may be calculated to be

$$\theta_1 = \tan^{-1} \left(\frac{h_2 - h_1}{d} \right) \quad (6.4)$$

$$\theta_2 = \tan^{-1} \left(\frac{h_2 + h_1}{d} \right) \quad (6.5)$$

Lastly, the time delays of arrival are given by $\tau_1 = \frac{r_1}{c}$ and $\tau_2 = \frac{r_2}{c}$, where c is the speed of propagation.

If we make the assumption that the antenna heights $h_1, h_2 \ll r_1, r_2$, we can use a binomial expansion on Eqs. (6.1) and (6.2) to simplify the distance expressions.

$$r_1 = \sqrt{d^2 + (h_2 - h_1)^2} \approx d + \frac{(h_2 - h_1)^2}{2d} \quad (6.6)$$

$$r_2 = \sqrt{d^2 + (h_2 + h_1)^2} \approx d + \frac{(h_2 + h_1)^2}{2d} \quad (6.7)$$

If we additionally assume that $r_1 \approx r_2$, we can substitute Eqs. (6.6) and (6.7) into Eq. (6.3) to get

$$\begin{aligned} E_{rs} &= \frac{E_0 e^{-jk r_1}}{r_1} \left[1 + \text{Re} e^{-jk(r_2 - r_1)} \right] \\ &= \frac{E_0 e^{-jk r_1}}{r_1} \left[1 + |R| e^{-j \left(k \frac{2h_1 h_2}{d} - \psi \right)} \right] \\ &= \frac{E_0 e^{-jk r_1}}{r_1} \left[1 + |R| \left(\cos \left(k \frac{2h_1 h_2}{d} - \psi \right) - j \sin \left(k \frac{2h_1 h_2}{d} - \psi \right) \right) \right] \end{aligned} \quad (6.8)$$

Equation (6.8) is in phasor form. We can transform this phasor back into the time domain by finding $\text{Re}\{E_{rs} e^{j\omega t}\}$. Through the use of Euler identities and some manipulation, we can write the time domain expression as

$$\begin{aligned} E_r(t) &= \frac{E_0}{r_1} \left[\left(1 + |R| \cos \left(k \frac{2h_1 h_2}{d} - \psi \right) \right) \cos(\omega t - k r_1) \right. \\ &\quad \left. + |R| \sin \left(k \frac{2h_1 h_2}{d} - \psi \right) \sin(\omega t - k r_1) \right] \end{aligned} \quad (6.9)$$

This solution is composed of two quadrature sinusoidal signals harmonically interacting with each other. If $R = 0$, Eq. (6.9) reverts to the direct path solution. Equation (6.9) is of the general form

$$X \cos(\omega t - kr_1) + Y \sin(\omega t - kr_1) = A \cos(\omega t - kr_1 + \phi) \quad (6.10)$$

where $X = \frac{E_0}{r_1} (1 + |R| \cos(k \frac{2h_1 h_2}{d} - \psi))$

$Y = \frac{E_0}{r_1} |R| \sin(k \frac{2h_1 h_2}{d} - \psi)$

$A = \sqrt{X^2 + Y^2}$ = signal envelope

$\phi = \tan^{-1} \left(\frac{Y}{X} \right)$ = signal phase

Thus, the envelope and phase of Eq. (6.9) is given by

$$A = \frac{E_0}{r_1} \sqrt{\left(1 + |R| \cos\left(\frac{2kh_1 h_2}{d} - \psi\right) \right)^2 + \left(|R| \sin\left(\frac{2kh_1 h_2}{d} - \psi\right) \right)^2} \quad (6.11)$$

$$\phi = \tan^{-1} \left(\frac{\frac{E_0}{r_1} |R| \sin\left(k \frac{2h_1 h_2}{d} - \psi\right)}{\frac{E_0}{r_1} \left(1 + |R| \cos\left(k \frac{2h_1 h_2}{d} - \psi\right) \right)} \right) \quad (6.12)$$

where $\frac{E_0}{r_1}$ = direct path amplitude.

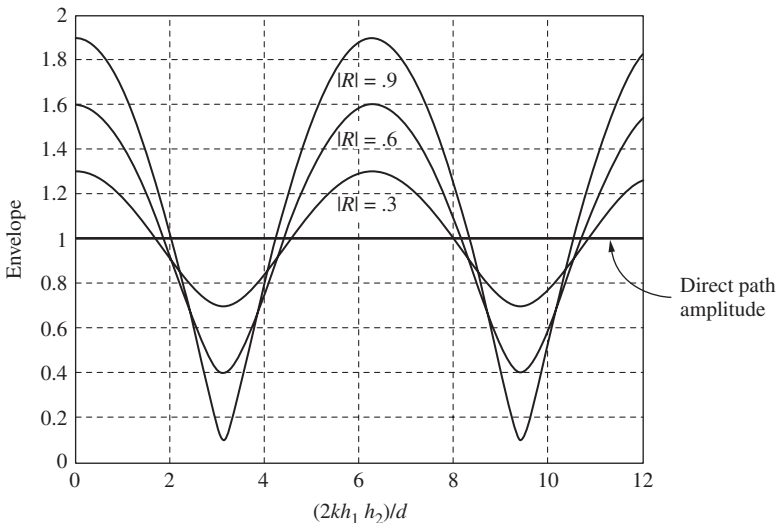


Figure 6.2 Envelope of direct and reflected signals over flat earth.

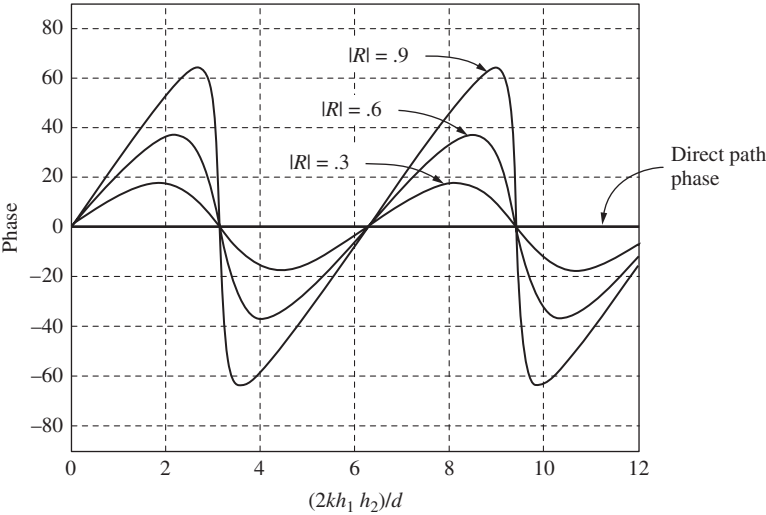


Figure 6.3 Phase of direct and reflected signals over flat earth.

Example 6.1 Assuming the following values: $|R| = .3, .6, .9, \psi = 0, E_0/r_1 = 1, h_1 = 5 \text{ m}, h_2 = 20 \text{ m}, d = 100 \text{ m}$. What are the time delays τ_1 and τ_2 ? Also, plot a family of curves for the envelope and the phase vs. $(2kh_1h_2)/d$

Solution Using Eqs. (6.6) and (6.7) we can calculate the distances $r_1 = 101.12 \text{ m}$ and $r_2 = 103.08 \text{ m}$. The corresponding time delays are $\tau_1 = .337 \mu\text{s}$ and $\tau_2 = .344 \mu\text{s}$. The plots of the envelope and phase are shown in Figs. (6.11) and (6.12)

Figures 6.2 and 6.3 demonstrate that if any one of the abscissa variables change, whether antenna heights (h_1, h_2), antenna distance (d), or the wavenumber k , the received signal will undergo significant interference effects. In general, the total received signal can fluctuate anywhere between zero and up to twice the amplitude of the direct path component. The phase angle can change anywhere between -90° and 90° . If antenna 1 represents a mobile phone in a moving vehicle, one can see how multipath can cause interference resulting in signal fading. Fading is defined as the *waxing* and *waning* of the received signal strength.

6.2 Multipath Propagation Mechanisms

The flat earth model discussed earlier is a deterministic model that is instructive in demonstrating the principles of basic interference. In the case of the flat earth model, we know the received signal voltage at all times. In typical multipath/channel scenarios, the distribution

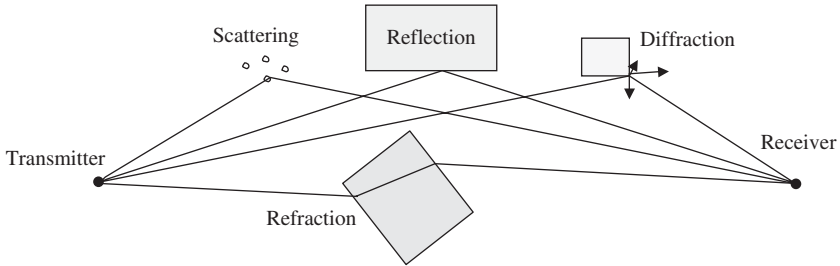


Figure 6.4 Different mechanisms creating multipath.

of large numbers of reflecting, diffracting, refracting, and scattering objects becomes random. In this case, numerous multiple paths can be created and it becomes extremely difficult to attempt to model the channel deterministically. Thus, one must revert to a statistical model for estimating signal and channel behavior.

Figure 6.4 shows a propagation channel with several candidate mechanisms for creating multiple propagation paths. Let us define these mechanisms.

- **Scattering:** Scattering occurs when the electromagnetic signal strikes objects that are much smaller than a wavelength. These objects could be water droplets, clouds, or insects for example. In the electromagnetic community this mechanism is often termed *Rayleigh scattering*. (This is not to be confused with Rayleigh fading although the two phenomena are interrelated).
- **Refraction:** Refraction occurs when an electromagnetic signal propagates through a structure. The propagation path is diverted because of the difference in the electrical properties of the medium. Boundary conditions help to determine the extent of the refraction.
- **Reflection:** Reflection occurs when an electromagnetic signal strikes a smooth surface at an angle and is reflected toward the receiver. The angle of reflection is equal to the angle of incidence under normal conditions.
- **Diffraction:** Diffraction occurs when the electromagnetic signal strikes an edge or corner of a structure that is large in terms of wavelength. The incident ray is diffracted in a cone of rays following Keller's laws of diffraction [10].

The various mechanisms such as scattering, refraction, reflection, and diffraction give rise to alternate propagation paths such that the received signal is a composite of numerous replicas all differing in phase, amplitude, and in time delay. Thus, the multipath signal amplitudes, phases, and time delays become random variables.

6.3 Propagation Channel Basics

In order to better understand the performance of a wireless signal propagating in a typical outdoor environment, it is necessary to define some terms. These terms are commonly used to describe properties or characteristics of the channel.

6.3.1 Fading

Fading is a term used to describe the fluctuations in a received signal as a result of multipath components. Several replicas of the signal arrive at the receiver, having traversed different propagation paths, adding constructively and destructively. The fading can be defined as fast or slow fading. Additionally, fading can be defined as *flat* or *frequency selective fading*.

Fast fading is propagation which is characterized by rapid fluctuations over very short distances. This fading is due to scattering from nearby objects and thus is termed *small-scale fading*. Typically fast fading can be observed up to half-wavelength distances. When there is no direct path (line-of-sight), a Rayleigh distribution tends to best fit this fading scenario, thus fast fading is sometimes referred to as *Rayleigh fading*. When there is a direct path or a dominant path, fast fading can be modeled with a Rician distribution.

Slow fading is propagation which is characterized by slow variations in the mean value of the signal. This fading is due to scattering from the more distant and larger objects and thus is termed *large-scale fading*. Typically slow fading is the trend in signal amplitude as the mobile user travels over large distances relative to a wavelength. The slow fading mean value is generally found by averaging the signal over 10 to 30 wavelengths [11]. A log-normal distribution tends to best fit this fading scenario, thus slow fading is sometimes referred to as *log-normal fading*. Figure 6.5 shows a superimposed plot of fast and slow fading.

Flat fading is when the frequency response of the channel is flat relative to the frequency of the transmit signal, that is, the channel bandwidth B_C is greater than the signal bandwidth B_S ($B_C > B_S$). Thus, the multipath characteristics of the channel preserve the signal quality at the receiver.

Frequency selective fading is when the channel bandwidth B_C is less than the signal bandwidth B_S ($B_C < B_S$). In this case, the multipath delays start to become a significant portion of the transmit signal time duration and dispersion occurs.

Fast fading is of particular interest to the electrical engineer because the resulting rapid fluctuations can cause severe problems in reliably maintaining communication.

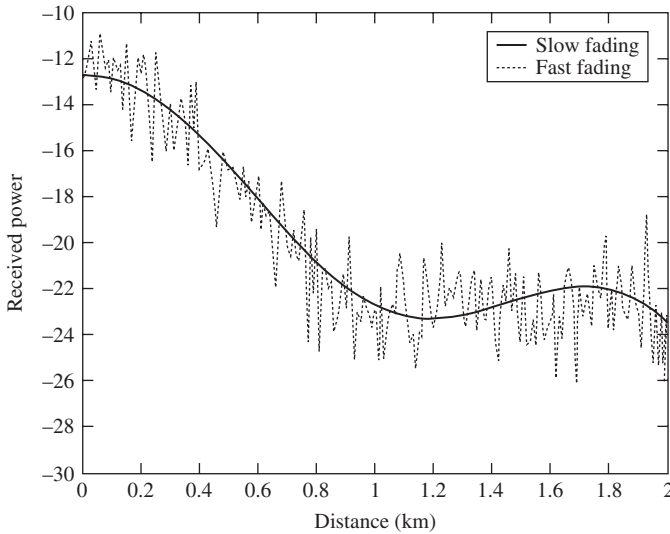


Figure 6.5 Slow and fast fading examples.

6.3.2 Fast fading modeling

Multipath with no direct path. Based upon the scenario as given in Fig. 6.4, we will assume that no direct path exists but that the entire received electric field is based upon multipath propagation. We can express the received voltage phasor as the sum of all the possible multipath component voltages within the receiver.

$$v_{rs} = \sum_{n=1}^N a_n e^{-j(kr_n - \alpha_n)} = \sum_{n=1}^N a_n e^{j\phi_n} \quad (6.13)$$

where a_n = random amplitude of the n th path

α_n = random phase associated with the n th path

r_n = length of n th path

$\phi_n = -kr_n + \alpha_n$

If we assume a large number of scattering structures N , which are randomly distributed, we can assume that the phases ϕ_n are uniformly distributed. We can express the time-domain version of the received voltage as

$$\begin{aligned} v_r &= \sum_{n=1}^N a_n \cos(\omega_0 t + \phi_n) \\ &= \sum_{n=1}^N a_n \cos(\phi_n) \cos(\omega_0 t) - \sum_{n=1}^N a_n \sin(\phi_n) \sin(\omega_0 t) \end{aligned} \quad (6.14)$$

We may further simplify Eq. (6.14) as we did in Eq. (6.10) using a simple trigonometric identity.

$$v_r = X \cos(\omega_0 t) - Y \sin(\omega_0 t) = r \cos(\omega_0 t + \phi) \quad (6.15)$$

$$\begin{aligned} \text{where } X &= \sum_{n=1}^N a_n \cos(\phi_n) \\ Y &= \sum_{n=1}^N a_n \sin(\phi_n) \\ r &= \sqrt{X^2 + Y^2} = \text{envelope} \\ \phi &= \tan^{-1} \left(\frac{Y}{X} \right) \end{aligned}$$

In the limit, as $N \rightarrow \infty$, the Central Limit Theorem dictates that the random variables X and Y will follow a Gaussian distribution with zero mean and standard deviation σ . The phase ϕ can also be modeled as a uniform distribution such that $p(\phi) = \frac{1}{2\pi}$ for $0 \leq \phi \leq 2\pi$. The envelope r is the result of a transformation of the random variables X and Y and can be shown to follow a Rayleigh distribution as given by Schwarz [12] or Papoulis [13]. The *Rayleigh probability density function* is defined as

$$p(r) = \frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}} \quad r \geq 0 \quad (6.16)$$

where σ^2 is the variance of the Gaussian random variables X or Y .

Figure 6.6 shows a Rayleigh distribution for two different standard deviations. Although the mean value of X or Y is zero, the mean value of the Rayleigh distribution is $\sigma\sqrt{\pi/2}$.

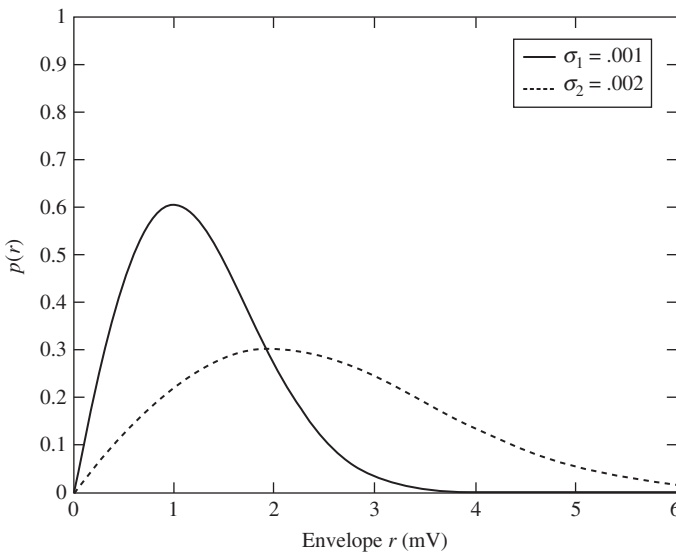


Figure 6.6 Rayleigh probability density.

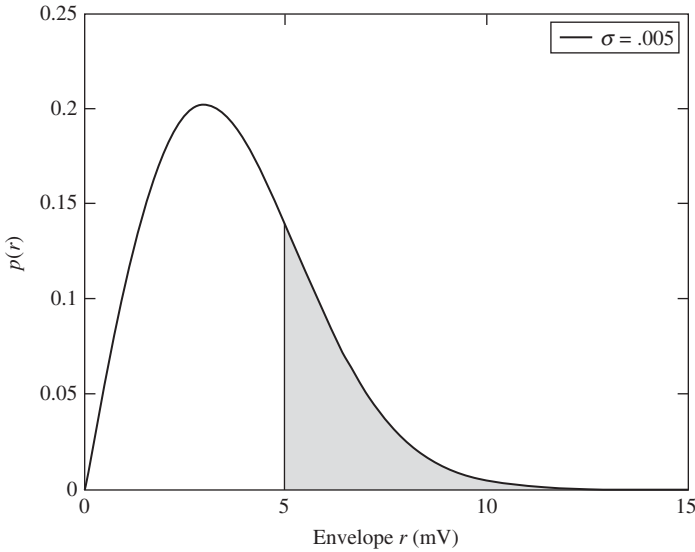


Figure 6.7 Rayleigh probability density with indicated threshold.

Example 6.2 For the Rayleigh-fading channel where $\sigma = .003$ V, what is the probability that the received voltage envelope will exceed a threshold of 5 mV?

Solution The probability of the envelope as exceeding 5 mV is given by

$$P(r \geq .005) = \int_{.005}^{\infty} \frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}} dr = .249$$

This can be shown as the shaded area under the curve in the Fig. 6.7.

It can be shown that if the envelope is Rayleigh distributed, the power p (watts) will have an exponential distribution [3, 13] (Also called an *Erlang distribution* with $n = 1$).

$$p(p) = \frac{1}{2\sigma^2} e^{-\frac{p}{2\sigma^2}} \quad p \geq 0 \quad (6.17)$$

The average value of the power is given by

$$\begin{aligned} E[p] &= p_0 = \int_0^{\infty} p \cdot p(p) dp \\ &= \int_0^{\infty} \frac{p}{2\sigma^2} e^{-\frac{p}{2\sigma^2}} dp = 2\sigma^2 \end{aligned} \quad (6.18)$$

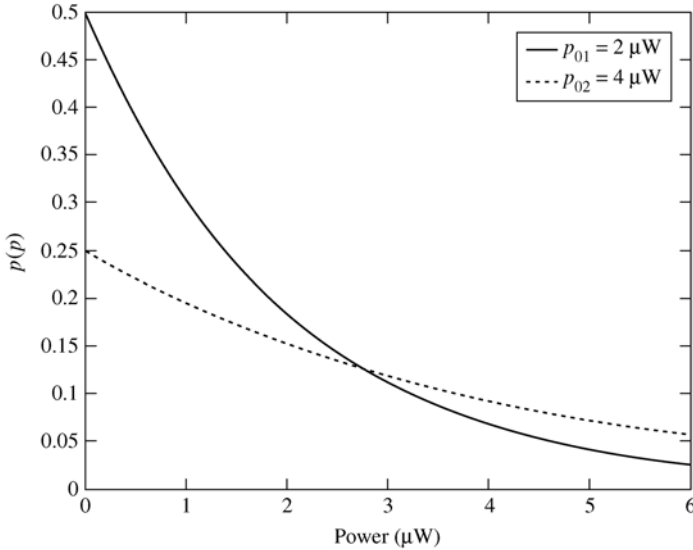


Figure 6.8 Exponential probability density.

Thus $p_0 = 2\sigma^2$ and can be substituted in Eq. (6.17). The power distribution is plotted in Fig. 6.8 for $p_0 = 2 \mu\text{W}$, $4 \mu\text{W}$.

The minimum detectable threshold power in a receiver is P_{th} . This power level is dictated by the receiver noise floor, noise figure, and detector thresholds. If the received power falls below the threshold, the receiver goes into “outage” because the backend signal-to-noise ratio is insufficient. The *outage probability* is the probability that the received power is too small for detection. This is given by

$$P(p \leq P_{\text{th}}) = \int_0^{P_{\text{th}}} \frac{1}{p_0} e^{-\frac{p}{p_0}} dp \quad (6.19)$$

Example 6.3 What is the outage probability of a Rayleigh channel if the average power is $2 \mu\text{W}$ and the threshold power is $1 \mu\text{W}$?

Solution The outage probability is given by

$$P(p \leq 1 \mu\text{W}) = \int_0^{1 \mu\text{W}} \frac{1}{2 \mu\text{W}} e^{-\frac{p}{2 \mu\text{W}}} dp = 0.393$$

This can be shown as the shaded area under the curve in the Fig. 6.9.

Multipath with direct path. Let us now consider that a direct path is present as indicated in Fig. 6.10. If a direct path is allowed in the

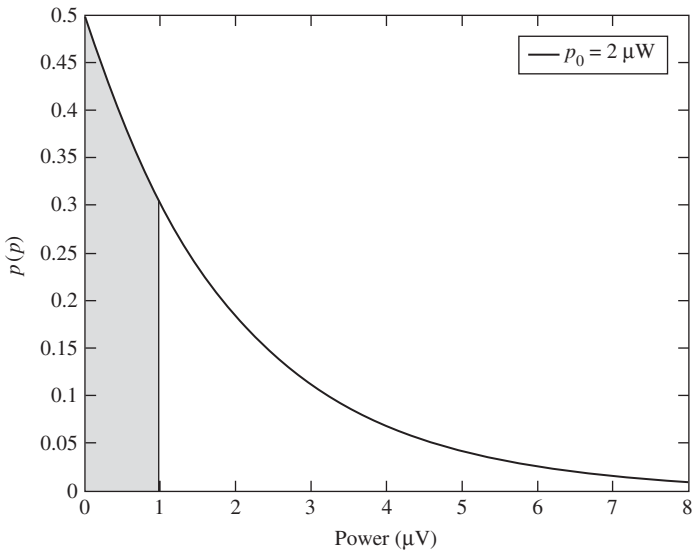


Figure 6.9 Outage probability shown as area under curve.

received voltage, we must modify Eqs. (6.14) and (6.15) by adding the direct path term with the direct path amplitude A (volts).

$$\begin{aligned} v_r &= A \cos(\omega_0 t) + \sum_{n=1}^N a_n \cos(\omega_0 t + \phi_n) \\ &= \left[A + \sum_{n=1}^N a_n \cos(\phi_n) \right] \cos(\omega_0 t) - \sum_{n=1}^N a_n \sin(\phi_n) \sin(\omega_0 t) \quad (6.20) \end{aligned}$$

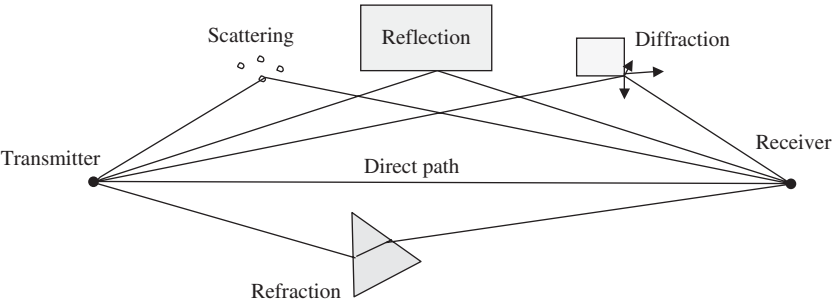


Figure 6.10 A multipath channel with a direct path.

Again, the envelope $r = \sqrt{X^2 + Y^2}$. We now must accordingly revise the random variables X and Y .

$$\begin{aligned} X &= A + \sum_{n=1}^N a_n \cos(\phi_n) \\ Y &= \sum_{n=1}^N a_n \sin(\phi_n) \end{aligned} \quad (6.21)$$

The random variable X is Gaussian with mean of A and standard deviation of σ . Random variable Y is Gaussian with zero mean and standard deviation of σ . The probability density function for the envelope is now a Rician distribution and is given by

$$p(r) = \frac{r}{\sigma^2} e^{-\frac{(r^2 + A^2)}{2\sigma^2}} I_0 \left(\frac{rA}{\sigma^2} \right) \quad r \geq 0 \quad A \geq 0 \quad (6.22)$$

where $I_0()$ = Modified Bessel function of first kind and zero-order.

We can characterize the Rician distribution by a parameter $K = A^2/(2\sigma^2)$. K is the direct signal power to multipath variance ratio. K is also called the *Rician factor*. We can also express K in dB as

$$K(\text{dB}) = 10 \log_{10} \left(\frac{A^2}{2\sigma^2} \right) \quad (6.23)$$

In the case where $A = 0$, the Rician distribution reverts to a Rayleigh distribution. A plot of the Rician distribution for three K values is shown in Fig. 6.11.

Example 6.4 For the Rician-fading channel where $\sigma = 3$ mV, the direct path amplitude $A = 5$ mV, what is the probability that the received voltage envelope will exceed a threshold of 5 mV?

Solution The probability of the envelope as exceeding 5 mV is given by

$$P(r \geq .005) = \int_{.005}^{\infty} \frac{r}{\sigma^2} e^{-\frac{(r^2 + A^2)}{2\sigma^2}} I_0 \left(\frac{rA}{\sigma^2} \right) dr = .627$$

This can be shown as the shaded area under the curve in Fig. 6.12.

As compared to Example 6.2 with no direct path, Example 6.4 demonstrates that the likelihood of having a detectable signal dramatically increases when the direct path is present.

Motion in a fast fading channel. In the above examples of *line-of-sight* (LOS) and *non-line-of-sight* (NLOS) propagation, it was assumed that

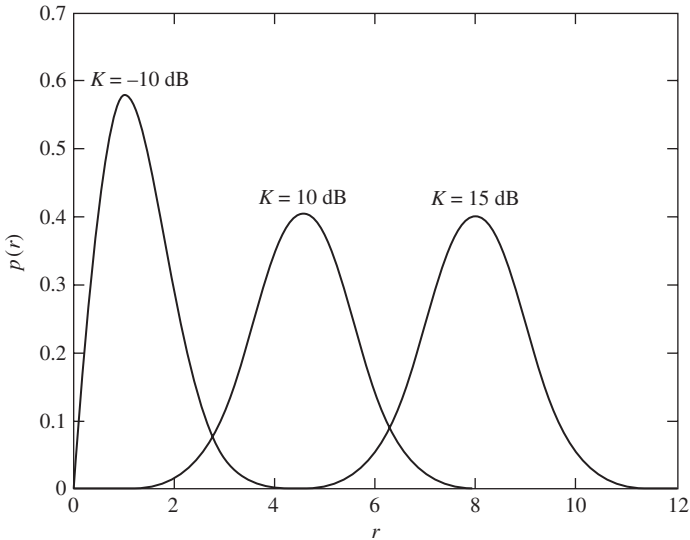


Figure 6.11 Rician distribution.

there was no motion in the channel. Motion will change the channel behavior by changing the location of the transmitter or receiver. In addition, motion introduces many discrete Doppler shifts in the received signal. Figure 6.13 shows a moving transmitter in a multipath environment with no direct path.

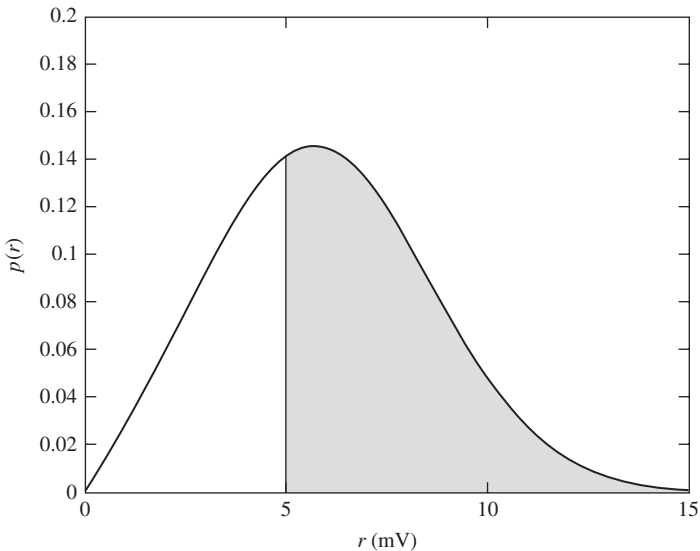


Figure 6.12 Rician probability density with indicated threshold.

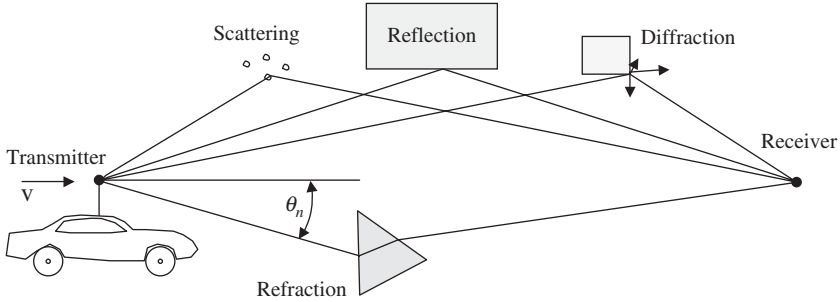


Figure 6.13 A multipath channel with no direct path.

As the vehicle moves at a constant velocity in Fig. 6.13, many factors change with time. The angles (θ_n) of each multipath signal are time dependent. Each multipath experiences a different Doppler shift because the angle of scattering with respect to the moving vehicle is different for each scattering object. Also, the overall phase shift (α_n) changes with time because the propagation delays are changing.

The maximum possible Doppler shift is given by

$$f_d = f_0 \frac{v}{c} \quad (6.24)$$

where f_d = Doppler frequency

f_0 = carrier frequency

v = vehicle velocity

c = speed of light

Since the direction of vehicle travel is at an angle θ_n with the n th multipath, the Doppler shift is modified accordingly. The Doppler shift for the n th path is given by

$$f_n = f_d \cos \theta_n = f_0 \frac{v}{c} \cos \theta_n \quad (6.25)$$

We can now rewrite Eq. (6.14) accounting for the Doppler frequency shifts f_n .

$$\begin{aligned} v_r &= \sum_{n=1}^N a_n \cos(2\pi f_n t + \phi_n) \cos(\omega_0 t) - \sum_{n=1}^N a_n \sin(2\pi f_n t + \phi_n) \sin(\omega_0 t) \\ &= \sum_{n=1}^N a_n \cos(2\pi f_d \cos(\theta_n) t + \phi_n) \cos(\omega_0 t) \\ &\quad - \sum_{n=1}^N a_n \sin(2\pi f_d \cos(\theta_n) t + \phi_n) \sin(\omega_0 t) \end{aligned} \quad (6.26)$$

We now have three random variables a_n , ϕ_n , and θ_n . The amplitude coefficients are Gaussian distributed whereas the phase coefficients are presumed to have a uniform distribution such that $0 \leq \phi_n$ and $\theta_n \leq 2\pi$. The envelope of v_r again has a Rayleigh distribution. The envelope r is given by

$$r = \sqrt{X^2 + Y^2} \tag{6.27}$$

where $X = \sum_{n=1}^N a_n \cos(2\pi f_d \cos(\theta_n)t + \phi_n)$ and
 $Y = \sum_{n=1}^N a_n \sin(2\pi f_d \cos(\theta_n)t + \phi_n)$

This model is called the *Clarke flat fading model* [7, 14].

Example 6.5 Use MATLAB to plot the envelope in Eq. (6.27) where the carrier frequency is 2 GHz, the vehicle velocity is 50 mph, the phase angles ϕ_n and θ_n are uniformly distributed, the coefficient a_n has a Gaussian distribution with zero mean and standard deviation of $\sigma = .001$. Let $N = 10$.

Solution We must first convert the velocity to meters/second. Therefore, 50 mph = 22.35 m/s. Thus, the maximum Doppler shift is

$$f_d = 2 \cdot 10^9 \cdot \frac{22.35}{3 \cdot 10^8} = 149 \text{ Hz}$$

Using the following short MATLAB program, we can plot the results as demonstrate in Fig. 6.14.

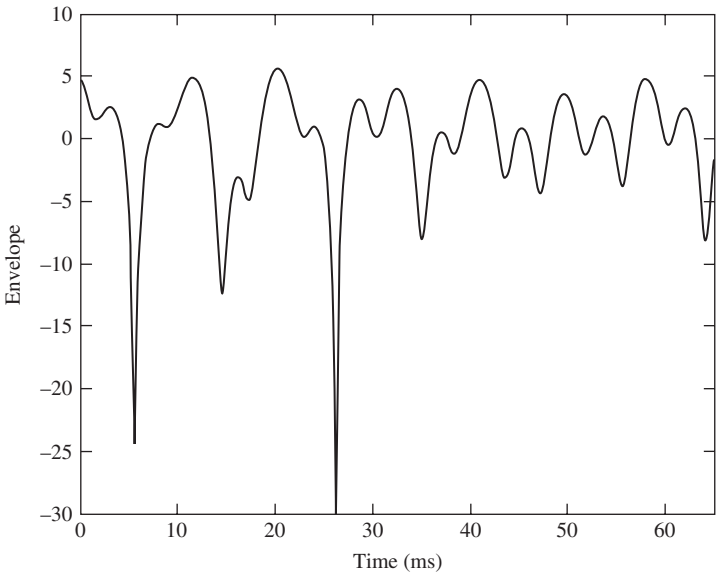


Figure 6.14 Doppler fading channel with $N = 10$ Paths.

% Fast fading with velocity Example 6.5

```

N = 10; % number of scatterers
a=.001*randn(N,1); % create Gaussian amplitude coefficients
th=rand(N,1)*2*pi; % create uniform phase angles
ph=rand(N,1)*2*pi;
fd=149; % Doppler
tmax = 10/fd; % Maximum time
omega=2*pi*fd;
t=[0:1000]*tmax/1000; % generate timeline
X=zeros(1,length(t));
Y=zeros(1,length(t));
for n=1:N % generate the sums for X and Y
X=X+a(n)*cos(omega*cos(th(n))*t+ph(n));
Y=Y+a(n)*sin(omega*cos(th(n))*t+ph(n));
end
r=sqrt(X.^2+Y.^2); % Calculate the Rayleigh envelope
rdb=20*log10(r); % Calculate the envelope in dB
figure;
plot(t*1000,rdb,'k') % plot
xlabel('time (ms)')
ylabel('envelope')
axis([0 65 -30 10])

```

Example 6.5 is representative of Doppler fading but makes assumptions that are not realistic. In the example, the scattering from objects is angularly dependent. Thus, the coefficients a_n will be a function of time. Additionally, the phase angles ϕ_n and θ_n change with time. The Clarke model can be modified to reflect the time dependence of a_n , ϕ_n , and θ_n .

If we assume a large number of paths, the uniform distribution of angles θ_n results in a sinusoidal variation in the Doppler frequencies f_n . This transformation of the random variable results in a Doppler power spectrum derived by Gans [15] and Jakes [16].

$$S_d(f) = \frac{\sigma^2}{\pi f_d \sqrt{1 - \left(\frac{f}{f_d}\right)^2}} \quad |f| \leq f_d \quad (6.28)$$

where $\sigma^2 = \sum_{n=1}^N E[a_n^2]$ = average power of signal.

A plot of the Doppler power spectrum is shown in Fig. 6.15.

As a test of the validity of Eq. (6.28) we can rerun the program of Example 6.5 by increasing the scatterers to 100 and padding the amplitude X with zeros. Again, the maximum Doppler shift is 149 Hz. After taking the fast Fourier transform, we can plot the results as seen in Fig. 6.16. The spectrum bears a similarity to the theoretical spectrum.

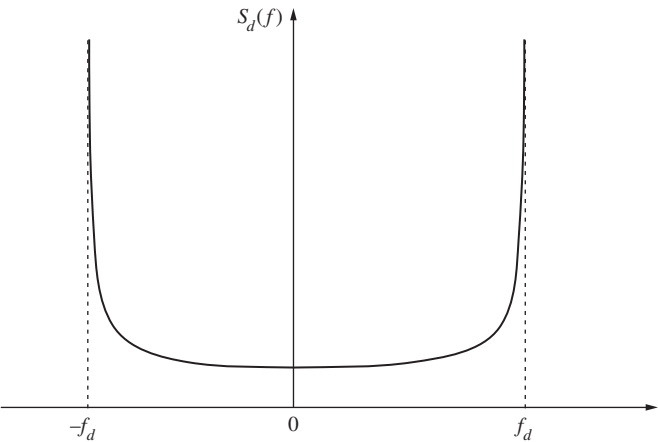


Figure 6.15 Doppler power density spectrum.

However, the difference lies in the fact that the theoretical spectrum assumes that the number of scatterers is large enough to apply the Central Limit Theorem. In that case, there is a true Gaussian distribution on the scattering amplitudes and a true uniform distribution on the angles.

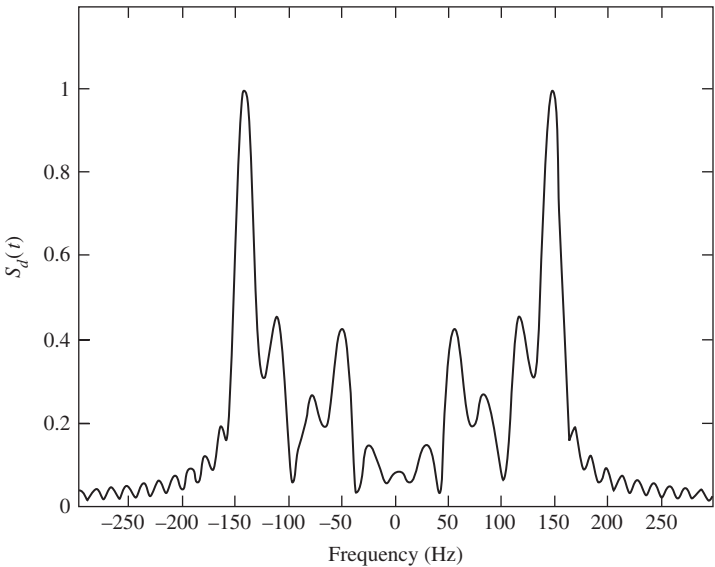


Figure 6.16 Doppler power density spectrum.

6.3.3 Channel impulse response

If we assume that the radio channel can be modeled as a linear filter then the characteristics of the channel can be modeled by finding the impulse response of the channel. Thus, all signal responses to the channel are dictated by the impulse response. The impulse response also gives an indication of the nature and number of the multiple paths. If it is assumed that the channel characteristics can change with time (i.e., the mobile user is moving), the channel impulse response will also be function of time. The generalized channel impulse response is given as follows:

$$h_c(t, \tau) = \sum_{n=1}^N a_n(t) e^{j\psi_n(t)} \delta(\tau - \tau_n(t)) \quad (6.29)$$

where $a_n(t)$ = time varying amplitude of path n

$\psi_n(t)$ = time varying phase of path n which can include the effects of Doppler

$\tau_n(t)$ = time varying delay of path n

An example of an impulse response magnitude, $|h_c(t, \tau)|$, is shown plotted in Fig. 6.17. The amplitudes generally diminish with increasing delay because the spherical spreading increases in proportion to the delay. Differences in reflection coefficients can cause an exception to this rule for similar path lengths because slightly more distant sources of scattering might have larger reflection coefficients. The discrete time-delayed impulses are sometimes called *fingers*, *taps*, or *returns*.

If we additionally assume that the channel is *wide sense stationary* (WSS) over small-scale times and distances, the impulse response will

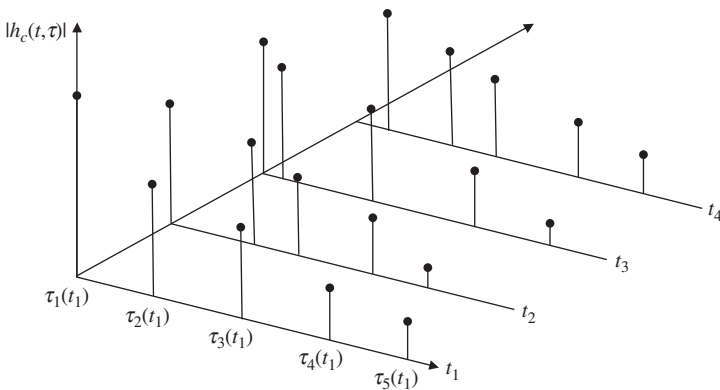


Figure 6.17 Channel impulse response at four instants in time.

remain stable for short time/fast fading. Thus, we can further simplify the impulse response to be approximated as

$$h_c(\tau) = \sum_{n=1}^N a_n e^{j\psi_n} \delta(\tau - \tau_n) \quad (6.30)$$

6.3.4 Power delay profile

In the case of small-scale fading, a channel metric can be defined to aid in understanding channel behavior. This metric is called the *power delay profile* (PDP). (The PDP is essentially identical to the *multipath intensity profile* (MIP) [17, 18]). The PDP can be defined as ([8, 9, 19])

$$P(\tau) = E[|h_c(t, \tau)|^2] = \sum_{n=1}^N P_n \delta(\tau - \tau_n) \quad (6.31)$$

where $P_n = \langle |a_n(t)|^2 \rangle = a_n^2$

$\tau_n = \langle \tau_n(t) \rangle$

$\langle x \rangle$ = estimate of the random value x

A typical power delay profile for an urban area is shown in Fig. 6.18 where the fingers are the result of scattering from buildings.

The PDP is partly defined by the trip time delays (τ_n). The nature and characteristics of the delays helps to define the expected channel performance. Thus, it is important to define some terms regarding trip delays. Several valuable texts define these statistics [3, 5, 6, 8].

- **First Arrival Delay (τ_A):** This is measured as the delay of the earliest arriving signal. The earliest arriving signal is either the shortest multipath or the direct path if a direct path is present. All other delays can be measured relative to τ_A . Channel analysis can be simplified by defining the first arrival delay as zero (i.e., $\tau_A = 0$).
- **Excess Delay:** This is the additional delay of any received signal relative to the first arrival delay τ_A . Typically, all delays are defined as excess delays.

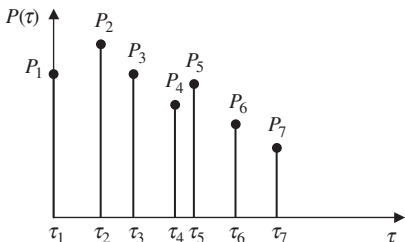


Figure 6.18 Typical urban power delay profile.

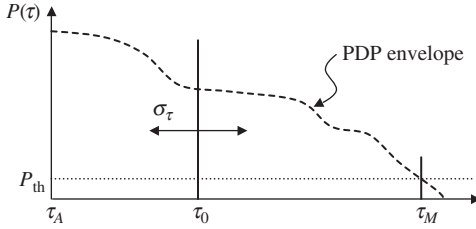


Figure 6.19 Depiction of PDP delay statistics.

- **Maximum Excess Delay (τ_M):** This is the maximum excess delay where the PDP is above a specified threshold. Thus, $P(\tau_M) = P_{th}$ (dB).
- **Mean Excess Delay (τ_0):** The mean value or first moment of all excess delays

$$\tau_0 = \frac{\sum_{n=1}^N P_n \tau_n}{\sum_{n=1}^N P_n} = \frac{\sum_{n=1}^N P_n \tau_n}{P_T} \quad (6.32)$$

where $P_T = \sum_{n=1}^N P_n$ = multipath power gain.

- **RMS Delay Spread (σ_τ):** This is the standard deviation for all excess delays.

$$\sigma_\tau = \sqrt{\langle \tau^2 \rangle - \tau_0^2} = \sqrt{\frac{\sum_{n=1}^N P_n \tau_n^2}{P_T} - \tau_0^2} \quad (6.33)$$

Figure 6.19 demonstrates the various terms defined earlier.

Example 6.6 Calculate the multipath power gain (P_T), the mean excess delay (τ_0), and the RMS delay spread (σ_τ) for the PDP given in the Fig. 6.20.

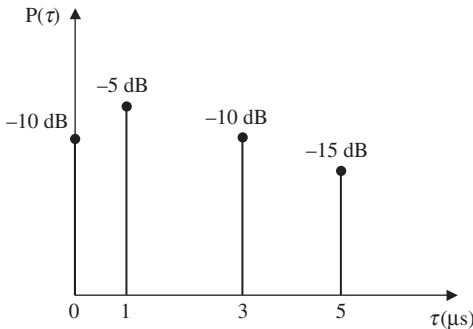


Figure 6.20 Power delay profile.

Solution First we must convert all powers to a linear scale. Thus, $P_1 = .1$, $P_2 = .32$, $P_3 = .1$, $P_4 = .032$. The multipath power gain is

$$P_T = \sum_{n=1}^N P_n = .552 \quad \text{or} \quad -2.58 \text{ dB.}$$

The mean excess delay is

$$\tau_0 = \frac{\sum_{n=1}^4 P_n \tau_n}{P_T} = \frac{.1 \times 0 + .32 \times 1 + .1 \times 3 + .032 \times 5}{.552} = 1.41 \mu\text{s}$$

The RMS delay spread is

$$\begin{aligned} \sigma_\tau &= \sqrt{\frac{\sum_{n=1}^N P_n \tau_n^2}{P_T} - \tau_0^2} \\ &= \sqrt{\frac{.1 \times 0^2 + .32 \times 1^2 + .1 \times 3^2 + .032 \times 5^2}{.552} - (1.41)^2} \\ &= 1.29 \mu\text{s} \end{aligned}$$

6.3.5 Prediction of power delay profiles

It is sometimes informative to model the power delay profiles in order to: (1) understand channel behavior, (2) evaluate the performance of equalizers, and (3) estimate *bit error rate* (BER) performance. Numerous measurements have been performed on indoor and outdoor channels and three relatively useful models have been proposed (Chuang [20], Feher [6]). The total received power is given by P_T . The three models are:

One-Sided Exponential Profile. This profile seems to most accurately describe both indoor and urban channels.

$$P(\tau) = \frac{P_T}{\sigma_\tau} e^{-\frac{\tau}{\sigma_\tau}} \quad \tau \geq 0 \quad (6.34)$$

Gaussian Profile

$$P(\tau) = \frac{P_T}{\sqrt{2\pi} \sigma_\tau} e^{-\frac{1}{2} \left(\frac{\tau}{\sigma_\tau}\right)^2} \quad (6.35)$$

Equal Amplitude Two-Ray Profile

$$P(\tau) = \frac{P_T}{2} [\delta(\tau) + \delta(\tau - 2\sigma_\tau)] \quad (6.36)$$

6.3.6 Power angular profile

The PDP and the path delays are instructive in helping one to understand the dispersive characteristics of the channel and to calculate the channel bandwidth. The PDP is especially relevant for *single-input single-output* (SISO) channels since one can view the impulse response as being for a SISO system. However, when an array is used at the receiver, the angles of arrival are of interest as well. Since the array has varying gain vs. angle-of-arrival ($G(\theta)$), it is instructive to also understand the statistics of the angles-of-arrival such as angular spread and mean angle-of-arrival (AOA). Every channel has angular statistics as well as delay statistics. Basic concepts in modeling the AOA have been addressed by Rappaport [9], Gans [15], Fulghum, Molnar, Duel-Hallen [21, 22], Boujemaa and Marcos [23], and Klein and Mohr [24].

This author will define the angular equivalent of the PDP as the *power angular profile* (PAP). Earlier references discuss the concept of a *power angle density* (PAD) [23] or a *power azimuth spectrum* (PAS) [22]. However, the PAD or the PAS is more analogous to the traditional *power spectral density* (PSD) rather than to the PDP used previously. The concept of a PAP immediately conveys angular impulse response information which assists in channel characterization. Thus the PAP is given as

$$P(\theta) = \sum_{n=1}^N P_n \delta(\theta - \theta_n) \quad (6.37)$$

Along with the PAP, we can define some metrics as indicators of the angular characteristics of the propagation paths.

- *Maximum Arrival Angle* (θ_M): This is the maximum angle relative to the boresight (θ_B) of the receive antenna array. The boresight angle is often the broadside direction for a linear array. The maximum angle is restricted such that $|\theta_M| - \theta_B \leq 180^\circ$.
- *Mean Arrival Angle* (θ_0): The mean value or first moment of all arrival angles

$$\theta_0 = \frac{\sum_{n=1}^N P_n \theta_n}{\sum_{n=1}^N P_n} = \frac{\sum_{n=1}^N P_n \theta_n}{P_T} \quad (6.38)$$

where $P_T = \sum_{n=1}^N P_n$ = multipath power gain.

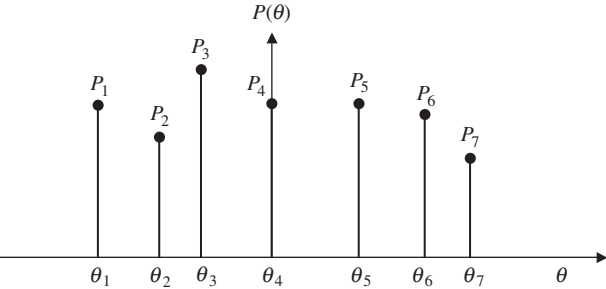


Figure 6.21 Power angular profile.

■ *RMS Angular Spread (σ_θ)*: This is the standard deviation for all arrival angles

$$\sigma_\theta = \sqrt{\frac{\sum_{n=1}^N P_n \theta_n^2}{P_T} - \theta_0^2} \tag{6.39}$$

Figure 6.21 depicts a representative PAP.

Example 6.7 Calculate the multipath power gain (P_T), the mean arrival angle (θ_0), and the RMS angular spread (σ_θ) for the PAP given in the Fig. 6.22

Solution First we must convert all powers to a linear scale. Thus, $P_1 = .1$, $P_2 = .32$, $P_3 = .1$, $P_4 = .032$. The multipath power gain is

$$P_T = \sum_{n=1}^N P_n = .552 \text{ W}$$

The mean arrival angle is

$$\begin{aligned} \theta_0 &= \frac{\sum_{n=1}^4 P_n \theta_n}{P_T} = \frac{.1 \times (-80) + .32 \times (-45) + .1 \times (40) + .032 \times (60)}{.552} \\ &= -29.86^\circ \end{aligned}$$

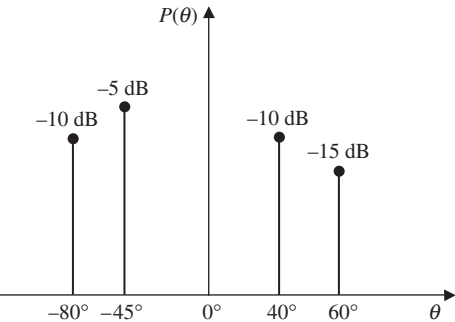


Figure 6.22 Power angular profile.

The RMS angular spread is

$$\begin{aligned}\sigma_\theta &= \sqrt{\frac{\sum_{n=1}^N P_n \theta_n^2}{P_T} - \theta_0^2} \\ &= \sqrt{\frac{.1 \times (-80)^2 + .32 \times (-45)^2 + .1 \times (40)^2 + .032 \times (60)^2}{.552} - (-29.86)^2} \\ &= 44^\circ\end{aligned}$$

An alternative approach for defining angular spread is given in Rappaport [9] where the angular spread is not found by defining first and second moments but rather the angular spread is determined through the use of a Fourier transform. This is akin to the use of the moment generating function in stochastic processes. The following description is slightly modified from the original Rappaport definition. We must first find the complex Fourier transform of the PAP. Thus

$$F_k = \int_0^{2\pi} P(\theta) e^{-jk\theta} d\theta \quad (6.40)$$

where $F_k = k$ th complex Fourier coefficient.

The angular spread is now defined as

$$\sigma_\theta = \theta_{\text{width}} \sqrt{1 - \frac{|F_1|^2}{F_0^2}} \quad (6.41)$$

where θ_{width} is the angular width of the PAP.

If indeed we use the power angular profile defined in Eq. (6.34), we can use the sifting property of the delta function and can find the Fourier coefficients.

$$F_k = \int_0^{2\pi} \sum_{n=1}^N P_n \delta(\theta - \theta_n) e^{-jk\theta} d\theta = \sum_{n=1}^N P_n e^{-jk\theta_n} \quad (6.42)$$

where

$$\begin{aligned}F_0 &= \sum_{n=1}^N P_n = P_T \\ F_1 &= \sum_{n=1}^N P_n e^{-j\theta_n}\end{aligned}$$

Example 6.8 Repeat Example 6.7 solving for the angular spread using the Rappaport method.

Solution The total angular width in Example 6.7 is 140° . The value of F_0 is the same as the total power, thus $F_0 = P_T = .552$. The value of F_1 , found from Eq. (6.39), is

$$\begin{aligned} F_1 &= \sum_{n=1}^4 P_n e^{-j\theta_n} = .1e^{j80^\circ} + .32e^{j45^\circ} + .1e^{-j40^\circ} + .032e^{-j60^\circ} \\ &= .34 + j.23 \end{aligned}$$

Substituting into Eq. (6.41) we can find the angular spread thus

$$\sigma_\theta = 94.3^\circ$$

This angular spread is larger than the angular spread in Example 6.7 by almost a factor of 2.

6.3.7 Prediction of angular spread

There are numerous potential models to describe angular spread under various conditions. An excellent overview is given by Ertel et al. [25]. Additionally, derivations are given for a ring of scatterers and a disk of scatterers in Pedersen, Mogensen, and Fleury [26]. Several models are given next showing the angular distribution for a ring of scatterers, disk of scatterers, and an indoor distribution of scatterers.

Ring of Scatterers: Figure 6.23 shows the transmitter and receiver where the transmitter is circled by a ring of scatterers uniformly distributed about the ring.

If we assume that the scattering occurs from a ring of scatterers, at constant radius, surrounding the transmitting antenna, then it can be shown that the PAP can be modeled as

$$P(\theta) = \frac{P_T}{\pi \sqrt{\theta_M^2 - \theta^2}} \quad (6.43)$$

where θ_M is the maximum arrival angle and P_T the total power in all angular paths.

A plot of this distribution is shown in Fig. 6.24 for the example where $P_T = \pi$ and $\theta_M = 45^\circ$.

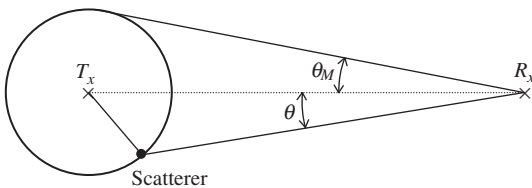


Figure 6.23 Ring of scatterers.

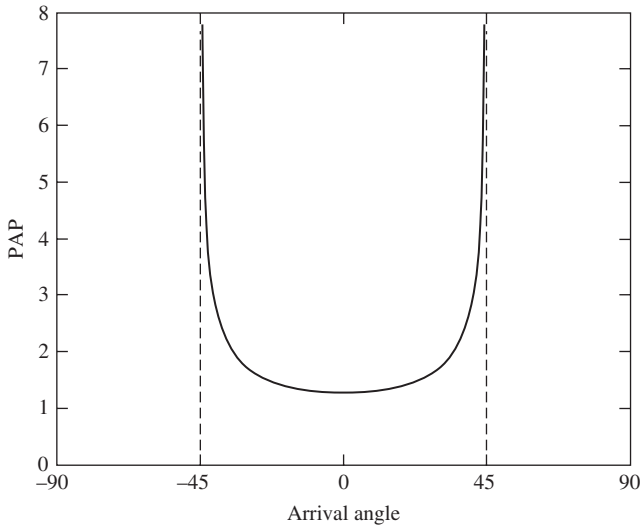


Figure 6.24 PAP for ring of scatterers.

Disk of Scatterers: Figure 6.25 shows the transmitter and receiver where the transmitter is enveloped by a uniform disk of scatterers.

The disk diameter is chosen to encompass the *circle of influence*. That is, the region whose radius encompasses the significant scatterers that create the largest multipath signals. Since the scatterers are uniformly distributed within the disk, the PAP is given by

$$P(\theta) = \frac{2P_T}{\pi \theta_M^2} \sqrt{\theta_M^2 - \theta^2} \tag{6.44}$$

A plot of this distribution is shown in Fig. 6.26 for the example where $P_T = \pi$ and $\theta_M = 45^\circ$.

Indoor Distribution: Figure 6.27 shows a typical transmitter and receiver indoors.

Extensive field tests have demonstrated that both indoor scattering and congested urban scattering can most closely be modeled by a

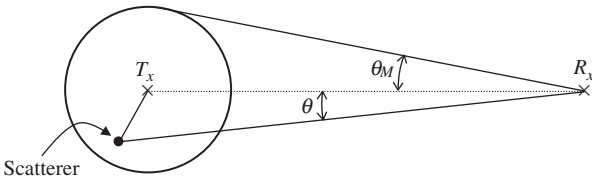


Figure 6.25 Disk of scatterers.

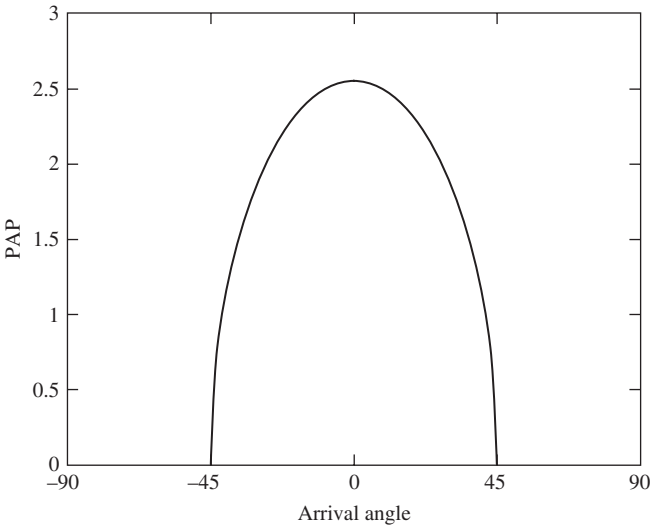


Figure 6.26 PAP for disk of scatterers.

Laplace distribution [22, 26, 27]. Thus, the PAP can be modeled as

$$P(\theta) = \frac{P_T}{\sqrt{2}\sigma_\theta} e^{-\left|\frac{\sqrt{2}\theta}{\sigma_\theta}\right|} \tag{6.45}$$

A plot of this distribution is shown in Fig. 6.28 for the example where $P_T = \pi$ and $\sigma_\theta = 30^\circ$.

The exponential distribution is intuitive because the largest received signal is the direct path or line-of-sight signal at 0° . Since the largest reflection coefficients are for near grazing angles, the smallest angles of arrival represent signals reflecting at grazing angles from the closest structures to the line-of-sight. As the angles further increase, the additional paths are more the result of diffraction and less the result of reflection. Diffraction coefficients are generally much smaller than reflection coefficients because diffraction produces an additional signal spreading. Wider arrival angles tend to represent higher order scattering mechanisms.

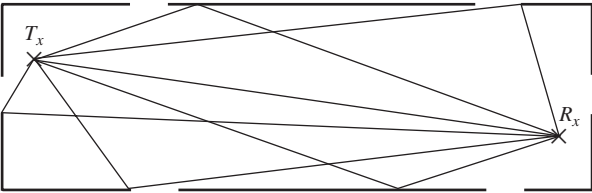


Figure 6.27 Indoor propagation.

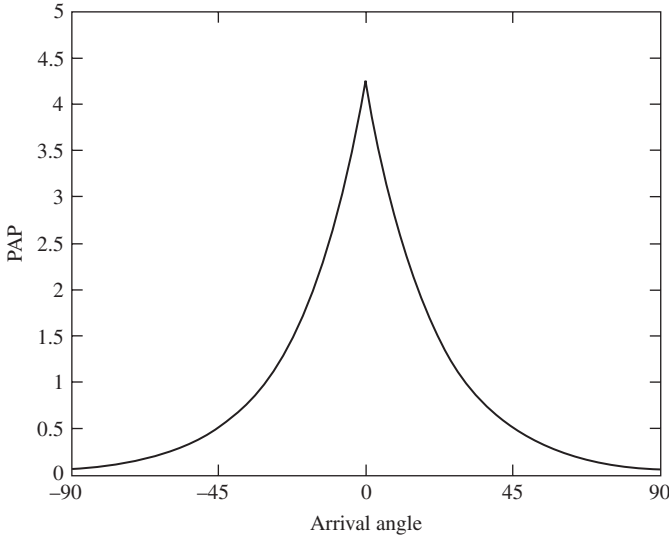


Figure 6.28 PAP for indoor scattering.

6.3.8 Power delay-angular profile

The next logical extension to the PDP and the PAP is the *power delay-angular profile* (PDAP). This can be viewed as the power profile resulting from the extended tap-delay-line method developed by Klein and Mohr [24], one that was further explored by Liberti and Rappaport [28]. This concept also can be derived from the work of Spencer et al. [27]. Thus, the PDAP is

$$P(\tau, \theta) = \sum_{n=1}^N P_n \delta(\tau - \tau_n) \delta(\theta - \theta_n) \quad (6.46)$$

We can combine the excess delays of Example 6.7 along with the arrival angles of Example 6.8 to generate a three-dimensional plot of the PDAP as shown in Fig. 6.29.

The extent of correlation between the time delays and angles of arrival will depend on the scattering obstacles in the channel. If most scatterers lie between the transmitter and the receiver, the correlation will be high. If the transmitter and receiver are surrounded by scatterers there may be little or no correlation.

6.3.9 Channel dispersion

From an electromagnetics perspective, dispersion normally occurs when, in a medium, the propagation velocities are frequency dependent. Thus, the higher frequencies of a transmitted signal will propagate at

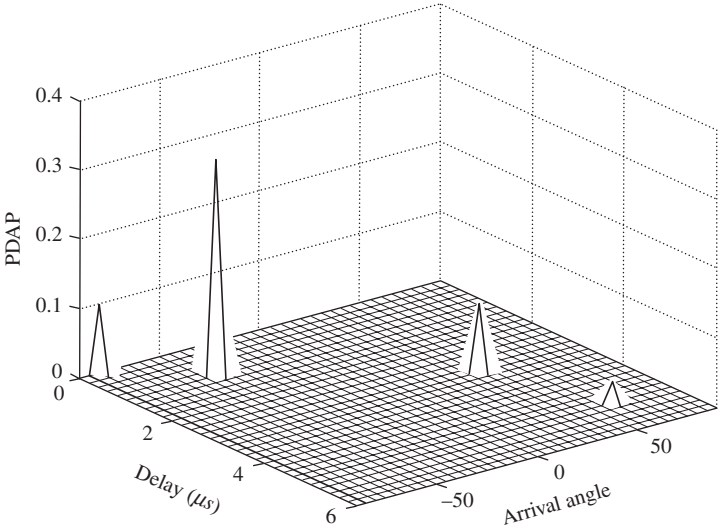


Figure 6.29 Power delay-angular profile.

different velocities than the lower frequencies. The high and low frequency components will produce different propagation delays even with no multipath present. This results in signal degradation at the receiver. However, a received signal can also be degraded by virtue of the fact that the time delays of the multipath components can become a reasonable fraction of the symbol period. Thus, the received signal is effectively degraded by time-delayed versions of itself. As the delay spread increases, the excess delays increase causing time dispersion. Figure 6.30 displays a direct path Gaussian pulse and the received signal for three increasing delay spreads (σ_1 , σ_2 , and σ_3). All signals are normalized. It is clear that the original pulse “broadens” as the delay spread increases. If the delay spread increases, this corresponds to a narrower channel bandwidth.

Since a channel can cause time-dispersion, it is necessary to define a channel bandwidth B_C , also referred to as a *coherence bandwidth*. This will help give an indication as to whether the channel bandwidth is sufficient to allow a signal to be transmitted with minimal dispersion. The channel bandwidth is described in Shankar [3] and Stein [29] and is approximately defined by

$$B_C = \frac{1}{5\sigma_\tau} \tag{6.47}$$

where σ_τ is the delay spread.

Therefore if the signal chip rate or bandwidth (B_S) is less than the channel bandwidth (B_C), the channel undergoes flat fading. If the signal

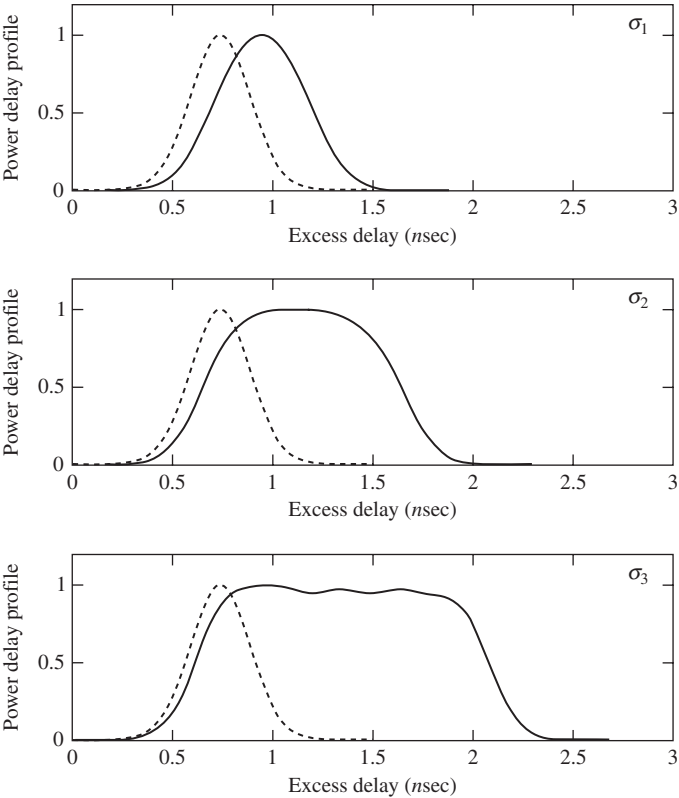


Figure 6.30 Dispersion caused by delay spread.

bandwidth is greater than the channel bandwidth, the channel will undergo frequency-selective fading. In this case dispersion occurs.

6.3.10 Slow fading modeling

Figure 6.5 demonstrates the trend of slow fading (also known as *shadow fading*). This represents the average, about which fast fading occurs. Instead of a single scattering mechanism, the transmitted signal can reflect, refract, diffract, and scatter multiple times before arriving at the receiver. Figure 6.31 demonstrates some of the candidate multipath mechanisms in the slow fading case along with coefficients representing each scattering location. It is assumed that no line-of-sight path exists.

The received signal can be represented as a sum of all multipath terms.

$$v_r(t) = \sum_{n=1}^N a_n e^{j\phi_n t} \tag{6.48}$$

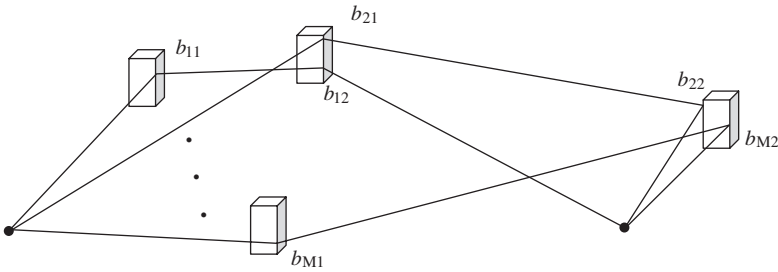


Figure 6.31 Slow fading with multiple scattering mechanisms.

The coefficients a_n represent the cascade product of each reflection or diffraction coefficient along path n . Thus, we can write a separate expression for the amplitude coefficients as

$$a_n = \prod_{m=1}^M b_{mn} \quad (6.49)$$

where b_{mn} = Rayleigh distributed random variables

M = number of scatterers along path n .

These multiple scattering events will affect the mean value of the received power. As was discussed earlier, the total power received (multipath power gain) is the sum of the square of the coefficients a_n . However, the coefficients a_n are the consequence of the products of the b_{mn} . The logarithm of the power is the sum of these random variables. Using the Central Limit Theorem, this becomes a Gaussian (normal) distribution. Hence the name *log-normal*. Several references on the log-normal distribution are Shankar [3], Agrawal and Zheng [4], Saunders [8], Rappaport[9], Lee [30], and Suzuki [31]. The pdf of the power in dBm is given as

$$p(P) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(P-P_0)^2}{2\sigma^2}} \quad (6.50)$$

where P = power p in dBm = $10 \cdot \log_{10}(p)$

P_0 = average signal level in dBm

σ = standard deviation in dBm

Using a transformation of the random variable P , we can also express the pdf in terms of linear power p .

$$p(p) = \frac{1}{p\sqrt{2\pi}\sigma_0} e^{-\frac{\log_{10}^2(\frac{p}{P_0})}{2\sigma_0^2}} \quad (6.51)$$

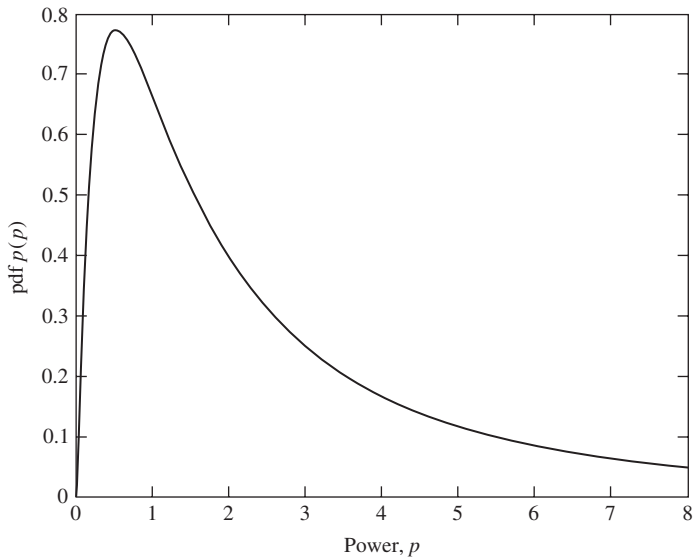


Figure 6.32 Log-normal density function.

where p = power in mW
 p_0 = average received signal level in mW
 $\sigma_0 = \frac{\log_{10}(\sigma)}{10}$

A typical plot of the log-normal distribution given in Eq. (6.51) is shown in Fig. 6.32.

6.4 Improving Signal Quality

One of the drawbacks of the typical multipath channel is the fact that the signal quality is degraded by Doppler spread and dispersion. The negative effects of dispersion were briefly discussed in Sec. 6.3.9. If the signal bandwidth is greater than the channel bandwidth, we have frequency-selective fading. The consequence is an increasing *intersymbol interference* (ISI) leading to an unacceptable BER performance. Thus data transmission rates are limited by the delay spread of the channel. Since we understand the nature of dispersion and Doppler spreading, we can devise methods to compensate in order to improve signal quality at the receiver. An excellent discussion of compensation techniques is found in Rappaport [9]. The three basic approaches to compensate for dispersion and Doppler spread are equalization, diversity, and channel coding.

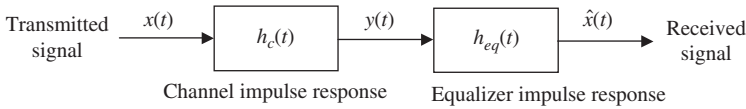


Figure 6.33 Communications system using an adaptive equalizer.

6.4.1 Equalization

Channel equalization is the act of reducing amplitude, frequency, time, and phase distortion created by a channel. The goal of the equalizer is to correct for the frequency selective fading effects in a channel where the signal bandwidth (B_S) exceeds the channel coherence bandwidth (B_C). The equalizer can be a signal processing algorithm that seeks to minimize the ISI. Unless the channel characteristics are fixed with time, the equalizer must be an adaptive equalizer which compensates as the channel characteristics change with time. The ultimate goal of the equalizer is to completely neutralize the negative effects of the channel. Figure 6.33 demonstrates the ideal equalizer. The goal of the equalizer impulse response is to negate the channel impulse response such that the receive signal is nearly identical to the transmitted signal.

The ideal equalizer frequency response which negates the channel influence would be given as

$$H_{eq}(f) = \frac{1}{H_c^*(-f)} \quad (6.52)$$

If the channel frequency response can be characterized by $H_c(f) = |H_c(f)|e^{j\phi_c(f)}$ then the equalizer frequency response would be given by $H_{eq}(f) = \frac{e^{j\phi_c(-f)}}{|H_c(-f)|}$ such that the product of the two filters is unity. The channel frequency response and equalizer response are shown in Figure 6.34.

Example 6.9 If the channel impulse response is given by $h_c(t) = .2\delta(t) + .3\delta(t - \tau)$, find the channel frequency response and the equalizer frequency response. Superimpose plots of the magnitude of $H_c(f)$ and $H_{eq}(f)$ for $0 \leq f\tau \leq 1.5$.

Solution The channel frequency response is the Fourier transform of the channel impulse response. Thus

$$\begin{aligned} H_c(f) &= \int_{-\infty}^{\infty} (.2\delta(t) + .3\delta(t - \tau))e^{-j2\pi ft} dt \\ &= .2 + .3e^{-j2\pi f\tau} \end{aligned}$$

The equalizer frequency response is given as

$$H_{eq}(f) = \frac{1}{H_c^*(-f)} = \frac{1}{.2 + .3e^{-j2\pi f\tau}} = \frac{.2 + .3e^{j2\pi f\tau}}{.13 + .12 \cos(2\pi f\tau)}$$

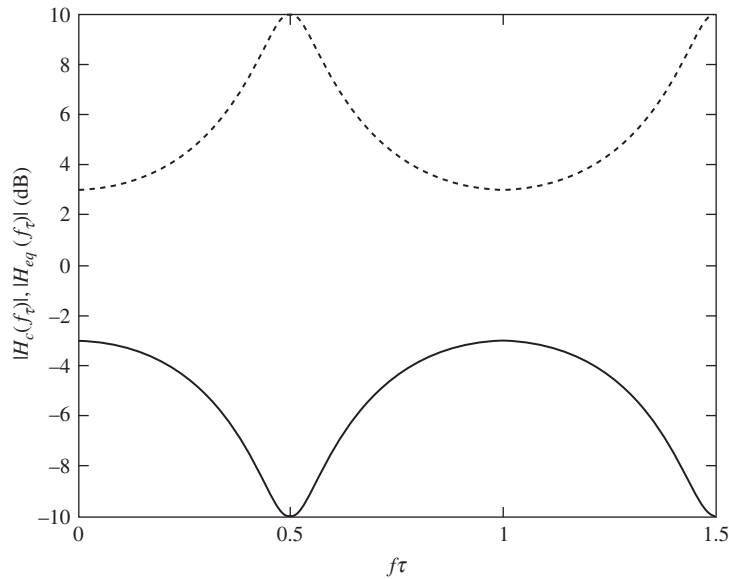


Figure 6.34 Channel and equalizer frequency response.

Since the channel characteristics are changing with time, it is necessary to design an adaptive equalizer that is able to respond to changing channel conditions. Figure 6.35 demonstrates a block diagram of an adaptive equalizer. The error signal is used as feedback to adjust the filter weights until the error is minimized.

The drawback of adaptive algorithms is that the algorithm must go through a *training* phase before it can *track* the received signal. The details of adaptive equalization are beyond the scope of this text but can be further explored in the Rappaport reference given previously.

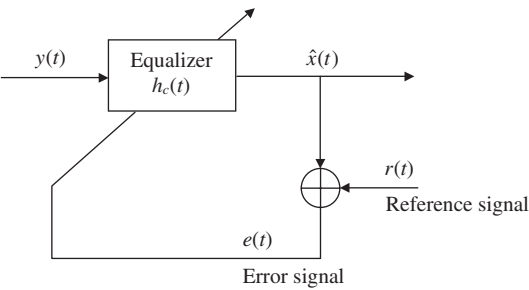


Figure 6.35 Adaptive equalizer.

6.4.2 Diversity

A second approach for mitigating the effects of fading is through diversity. Diversity means that a plurality of information is transmitted or received allowing one to reduce the depth and/or the duration of fades. This plurality can be achieved by having multiple receive antenna elements (antenna space diversity), different polarizations (polarization diversity), different transmit frequencies (frequency diversity), or different time characteristics such as Code division multiple access CDMA (time diversity). Thus, one can choose, amongst a collection of received signals, which signal provides the least amount of fading. This can be achieved by choosing perhaps the optimum antenna element, the optimum polarization, the optimum carrier frequency, or the optimum time diversity signal. Diversity has an advantage in that no adaptation or training is required in order to optimize the receiver.

Space diversity can be implemented through one of four basic approaches: selection diversity (select the largest signal from the antenna outputs), feedback diversity (scan all antenna outputs to find the first with a sufficient SNR for detection), maximal ratio combining (weight, co-phase, and sum all antenna outputs), and equal-gain combining (co-phase all received signals and combine with unity weights).

Polarization diversity is normally implemented through orthogonal polarizations. Orthogonal polarizations are uncorrelated and one can receive both polarizations with a dual polarized receive antenna. In addition, left and right hand circular polarizations can be used. One can select the polarization that maximizes the received signal-to-noise ratio.

Frequency diversity can be implemented by using multiple carrier frequencies. Presumably if a deep fade occurs at one frequency, the fade may be less pronounced at another frequency. It is suggested that the frequency separation be at least the channel coherence bandwidth (B_C) in order to ensure that the received signals are uncorrelated. Thus, the collection of possible transmit frequencies can be $f_0 \pm nB_C$ ($n = 0, 1, 2, \dots$).

Time diversity can be implemented by transmitting the same information multiple times where the time delay exceeds the coherence time ($T_c = 1/B_C$). Thus, the times of transmission can be $T_0 + nT_c$ ($n = 0, 1, 2, \dots$). Time diversity can also be implemented by modulating the transmission with CDMA pseudo-random codes. Typically these pn codes are uncorrelated if the different signal paths cause time delays exceeding a chip width. A RAKE receiver is a possible implementation of a time diversity scheme.

RAKE receiver. A channelized correlation receiver can be used where each channel attempts a correlation with the received signal based

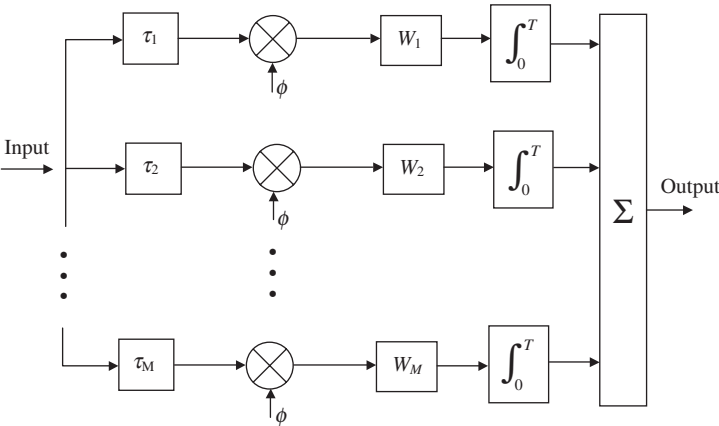


Figure 6.36 Structure of a RAKE receiver.

upon the anticipated path delays. One channel is assigned for each of the anticipated M strongest received components. The time delay τ_m is associated with the anticipated delay of the m th path. The correlating CDMA waveform is given by ϕ . Each channel can also have a weight W_n to allow the receiver to produce maximal ratio combining or equal-gain combining. Figure 6.36 demonstrates such a channelized receiver called a *RAKE receiver*. The RAKE receiver received its name based upon its similarity to a garden rake (Price and Green [32]).

As an example of a two channel RAKE receiver, we can transmit a 32-chip CDMA waveform of length $T = 2 \mu\text{s}$. Let there only be two paths to the receiver. The excess delay for path 1 is $0 \mu\text{s}$ and the excess delay for path 2 is 71.6 ns or 1.3 times the chip width ($\tau_{\text{chip}} = T/32 = 62.5 \text{ ns}$). The received waveforms are shown in Fig. 6.37. The correlation of the first waveform with itself is greater than four times the correlation with the delayed copy of itself. Thus, the first channel has a high correlation with the first arriving waveform and the second channel has a high correlation with the second arriving waveform. The first channel excess delay $\tau_1 = 0 \mu\text{s}$. The second channel excess delay $\tau_2 = 71.6 \text{ ns}$.

6.4.3 Channel coding

Because of the adverse fading effects of the channel, digital data can be corrupted at the receiver. Channel coding deliberately introduces redundancies into the data to allow for correction of the errors caused by dispersion. These redundant symbols can be used to detect errors and/or correct the errors in the corrupted data. Thus, the channel codes can fall into two types: *error detection codes* and *error correction codes*. The error detection codes are called *automatic repeat request (ARQ) codes*.

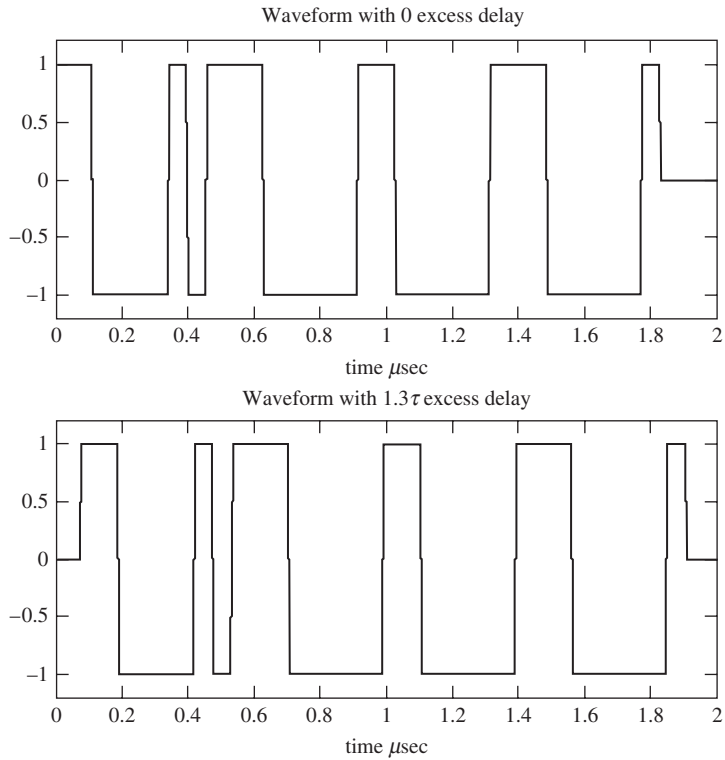


Figure 6.37 Channel 1 and 2 waveforms.

The error correction codes are *forward error correcting* (FEC) codes. The combination of both is called a *hybrid-ARQ code*. Two basic types of codes can be used to accomplish error detection and/or error correction. They are *block codes* and *convolutional codes*. *Turbo codes* are a type of convolutional code. An excellent description of these codes and channel coding schemes is given by Rappaport [9], Sklar [10], Proakis [19], and Parsons [33].

6.4.4 MIMO

MIMO stands for multiple-input multiple-output communication. This is the condition under which the transmit and receive antennas have multiple antenna elements. This is also referred to as *volume-to-volume* or as a *multiple-transmit multiple-receive* (MTMR) communications link. MIMO has applications in broadband wireless, WLAN, 3G, and other related systems. MIMO stands in contrast to *single-input single-output* (SISO) systems. Since MIMO involves multiple antennas, it can

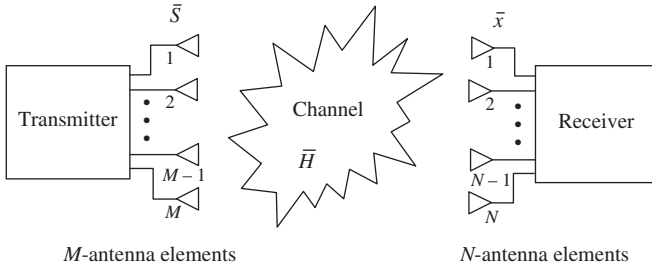


Figure 6.38 MIMO system.

be viewed as a space diversity approach to channel fading mitigation. Excellent sources for MIMO systems information can be found in the IEEE Journal on Selected Areas in Communications: MIMO Systems and Applications Parts 1 and 2 [34, 35], Vucetic and Yuan [36], Dig-gavi et al. [37], and Haykin and Moher [38]. A basic MIMO system is illustrated in Fig. 6.38 where $\bar{H}$ is the complex channel matrix relating M inputs to N outputs, $\bar{s}$ is the complex transmit vector, and $\bar{x}$ is the complex receive vector.

The goal of MIMO is to combine signals on both the transmit and receive ends such that the data rate is increased and/or the ISI and BER are decreased. As was mentioned previously, one diversity option for a SISO system is time diversity. MIMO allows the user to combine time diversity and space diversity. It can be said that MIMO is more than a space diversity option but is a space-time signal processing solution to channel fading. Multipath propagation has been viewed as a nuisance in SISO systems but it is actually an advantage for MIMO systems because the multipath information can be combined to produce a better received signal. Thus the goal of MIMO is not to mitigate channel fading but rather to take advantage of the fading process.

We may follow the development in Haykin and Moher [38] in order to understand the mathematical description of the MIMO system. Under flat-fading conditions, let us define the M -element complex transmit vector, created by the M -element transmit array, as

$$\bar{s} = [s_1 \quad s_2 \quad s_3 \quad \dots \quad s_M]^T \quad (6.53)$$

The elements s_m are phased according to the array configuration and are assumed to be coded waveforms in the form of a data stream. If we assume that the transmit vector elements have zero mean and variance σ_s^2 , the total transmit power is given by

$$P_t = M \cdot \sigma_s^2 \quad (6.54)$$

Each transmit antenna element, m , connects to a path (or paths) to the receive elements, n , creating a channel transfer function, h_{nm} . There are consequently $N \cdot M$ channel transfer functions connecting the transmit and receive array elements. Thus, we can define the $N \times M$ complex channel matrix as

$$\bar{H} = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1M} \\ h_{21} & h_{22} & \cdots & h_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ h_{N1} & h_{N2} & \cdots & h_{NM} \end{bmatrix} \quad (6.55)$$

Let us now define the N -element complex receive vector, created by the N -element receive array, as

$$\bar{x} = [x_1 \quad x_2 \quad x_3 \quad \dots \quad x_N]^T \quad (6.56)$$

Let us also define the N -element complex channel noise vector as

$$\bar{n} = [n_1 \quad n_2 \quad n_3 \quad \dots \quad n_N]^T \quad (6.57)$$

We can now describe the receive vector in matrix form as

$$\bar{x} = \bar{H} \cdot \bar{s} + \bar{n} \quad (6.58)$$

Assuming a Gaussian distribution for the transmitted signal, the channel, the received signal, and the noise, one can easily estimate the MIMO channel capacity. The correlation matrix for the transmitted signal is given by

$$\begin{aligned} \bar{R}_s &= E[\bar{s} \cdot \bar{s}^H] \\ &= \sigma_s^2 \bar{I}_M \end{aligned} \quad (6.59)$$

where σ_s^2 = signal variance

$\bar{I}_M = M \times M$ identity matrix.

Here $\bar{I}_M$ is the $M \times M$ identity matrix and $\bar{s}^H$ denotes the Hermitian transpose of $\bar{s}$. The correlation matrix for the noise is given by

$$\begin{aligned} \bar{R}_n &= E[\bar{n} \cdot \bar{n}^H] \\ &= \sigma_n^2 \bar{I}_N \end{aligned} \quad (6.60)$$

where σ_n^2 = noise variance

$\bar{I}_N = N \times N$ identity matrix.

The MIMO channel capacity is a random variable. Since the channel itself is random, we can define the ergodic (mean) capacity. If we

assume that the sources are uncorrelated and of equal power, the ergodic capacity is given as [39–41]

$$C_{EP} = E \left[\log_2 \left[\det \left(\bar{I}_N + \frac{\rho}{M} \bar{H} \cdot \bar{H}^H \right) \right] \right] \text{ bits/s/Hz} \quad (6.61)$$

where, the expectation is over the random channel matrix $\bar{H}$ and

$\det$ = determinant

C_{EP} = equal power capacity

ρ = SNR at each receive antenna = $\frac{P_t}{\sigma_n^2}$

One possible algorithm for use in a MIMO system is the V-BLAST algorithm developed at Bell Labs [42]. The V-BLAST algorithm is an improvement over its predecessor D-BLAST. (V-BLAST stands for Vertically layered blocking structure, Bell Laboratories Layered Space-Time). The V-BLAST algorithm demultiplexes a single data stream into M substreams which undergo a bit-to-symbol mapping. The mapped substreams are subsequently transmitted from the M transmit antennas. Thus, the total channel bandwidth used is a fraction of the original data stream bandwidth allowing for flat fading. The detection at the receiver can be performed by conventional adaptive beamforming. Each substream is considered to be the desired signal and all other substreams are deemed as interference and are therefore nulled. Thus, M simultaneous beams must be formed by the N -element receive array while nulling the unwanted substreams. The received substreams may then be multiplexed to recover the intended transmission.

A further exploration of MIMO concepts can be pursued in references [34–42].

References

1. Liberti, J.C., and T.S. Rappaport, *Smart Antennas for Wireless Communications: IS-95 and Third Generation CDMA Applications*, Prentice Hall, New York, 1999.
2. Bertoni, H.L., *Radio Propagation for Modern Wireless Systems*, Prentice Hall, New York, 2000.
3. Shankar, P.M., *Introduction to Wireless Systems*, Wiley, New York, 2002.
4. Agrawal, D.P., and Q.A. Zeng, *Introduction to Wireless and Mobile Systems*, Thomson Brooks/Cole, Toronto, Canada, 2003.
5. Sarkar, T.K., M.C. Wicks, M. Salazar-Palma, et al., *Smart Antennas*, IEEE Press & Wiley Interscience, New York, 2003.
6. Feher, K., *Wireless Digital Communications: Modulation and Spread Spectrum Applications*, Prentice Hall, New York, 1995.
7. Haykin, S. and M. Moher, *Modern Wireless Communications*, Prentice Hall, New York, 2005.
8. Saunders, S.R., *Antennas and Propagation for Wireless Communication Systems*, Wiley, New York, 1999.
9. Rappaport, T.S., *Wireless Communications: Principles and Practice*, 2d ed., Prentice Hall, New York, 2002.
10. Keller, J.B., "Geometrical Theory of Diffraction," *J. Opt. Soc. Amer.*, Vol. 52, pp. 116–130, 1962.

11. Sklar, B., *Digital Communications: Fundamentals and Applications*, 2d ed., Prentice Hall, New York, 2001.
12. Schwartz, M., *Information Transmission, Modulation, and Noise*, 4th ed., McGraw-Hill, New York, 1990.
13. Papoulis, A., *Probability, Random Variables, and Stochastic Processes*, 2d ed., McGraw-Hill, New York, 1984.
14. Clarke, R.H., "A Statistical Theory of Mobile-Radio Reception," *Bell Syst. Tech. J.*, Vol. 47, pp. 957–1000, 1968.
15. Gans, M.J., "A Power Spectral Theory of Propagation in the Mobile Radio Environment," *IEEE Trans. Veh. Technol.*, Vol. VT-21, No. 1, pp. 27–38, Feb. 1972.
16. Jakes, W.C., (ed.), *Microwave Mobile Communications*, Wiley, New York, 1974.
17. Ghassemzadeh, S.S., L.J. Greenstein, T. Sveinsson, et al., "A Multipath Intensity Profile Model for Residential Environments," *IEEE Wireless Communications and Networking*, Vol.1, pp.150–155, March 2003.
18. Proakis, J.G., *Digital Communications*, 2d ed., McGraw-Hill, New York, 1989.
19. Wesolowski, K., *Mobile Communication Systems*, Wiley, New York, 2004.
20. Chuang, J., "The Effects of Time Delay Spread on Portable Radio Communications Channels with Digital Modulation," *IEEE Journal on Selected Areas in Communications*, Vol. SAC-5, No. 5, June 1987.
21. Fulghum, T., and K. Molnar, "The Jakes Fading Model Incorporating Angular Spread for a Disk of Scatterers," *IEEE 48th Vehicular Technology Conference*, Vol.1, pp. 489–493, 18–21 May 1998.
22. Fulghum, T., K. Molnar, A. Duel-Hallen, "The Jakes Fading Model for Antenna Arrays Incorporating Azimuth Spread," *IEEE Transactions on Vehicular Technology*, Vol. 51, No. 5, pp. 968–977, Sept. 2002.
23. Boujemaa, H., and S. Marcos, "Joint Estimation of Direction of Arrival and Angular Spread Using the Knowledge of the Power Angle Density," *Personal, Indoor and Mobile Radio Communications, 13th IEEE International Symposium*, Vol. 4, pp. 1517–1521, 15–18 Sept. 2002.
24. Klein, A., and W. Mohr, "A Statistical Wideband Mobile Radio Channel Model Including the Directions-of-Arrival," *Spread Spectrum Techniques and Applications Proceedings, IEEE 4th International Symposium on*, Vol. 1, Sept. 1996.
25. Ertel, R., P. Cardieri, K. Sowerby, et al., "Overview of Spatial Channel Models for Antenna Array Communication Systems," *IEEE Personal Communications*, pp. 10–22, Feb 1998.
26. Pedersen, K., P. Mogensen, and B. Fleury, "A Stochastic Model of the Temporal and Azimuthal Dispersion Seen at the Base Station in Outdoor Propagation Environments," *IEEE Transactions on Vehicular Technology*, Vol. 49, No. 2, March 2000.
27. Spencer, Q., M. Rice, B. Jeffs, et al., "A Statistical Model for Angle of Arrival in Indoor Multipath Propagation," *Vehicular Technology Conference, 1997 IEEE 47th*, Vol. 3, pp. 1415–1419, 4–7 May 1997.
28. Liberti, J.C., and T.S. Rappaport, "A Geometrically Based Model for Line-of-Sight Multipath Radio Channels," *IEEE 46th Vehicular Technology Conference*, 1996. 'Mobile Technology for the Human Race', Vol. 2, pp. 844–848, 28 April–1 May 1996.
29. Stein, S., "Fading Channel Issues in System Engineering," *IEEE Journal on Selected Areas in Communications*, Vol. SAC-5, No. 2, pp. 68–89, Feb 1987.
30. Lee, W.C.Y., "Estimate of Local Average Power of a Mobile Radio Signal," *IEEE Transactions Vehicular Technology*, Vol. 29, pp. 93–104, May 1980.
31. Suzuki, H., "A Statistical Model for Urban radio Propagation," *IEEE Transactions on Communications*, Vol. COM-25, No. 7, 1977.
32. Price, R., and P. Green, "A Communication Technique for Multipath Channels," *Proceedings of IRE*, Vol. 46, pp. 555–570, March 1958.
33. Parsons, J., *The Mobile Radio Propagation Channel*, 2d ed., Wiley, New York, 2000.
34. "MIMO Systems and Applications Part 1," *IEEE Journal on Selected Areas in Communications*, Vol. 21, No. 3, April 2003.
35. "MIMO Systems and Applications Part 2," *IEEE Journal on Selected Areas in Communications*, Vol. 21, No. 5, June 2003.
36. Vucetic, B., and J. Yuan, *Space-Time Coding*, Wiley, New York, 2003.

37. Diggavi, S., N. Al-Dhahir, A. Stamoulis, et al., "Great Expectations: The Value of Spatial Diversity in Wireless Networks," *Proceedings of the IEEE*, Vol. 92, No. 2, Feb. 2004.
38. Haykin, S., and M. Moher, *Modern Wireless Communications*, Prentice Hall, New York, 2005.
39. Foschini, G., and M. Gans, "On Limits of Wireless Communications in a Fading Environment when Using Multiple Antennas," *Wireless Personal Communications* 6, pp. 311–335, 1998.
40. Gesbert, D., M. Shafi, D. Shiu, et al., "From Theory to Practice: An Overview of MIMO Space-Time Coded Wireless Systems," *Journal on Selected Areas in Communication*, Vol. 21, No. 3, April 2003.
41. Telatar, E., "Capacity of Multiantenna Gaussian Channels," *AT&T Bell Labs Technology Journal*, pp. 41–59, 1996.
42. Golden, G., C. Foschini, R. Valenzuela, et al., "Detection Algorithm and Initial Laboratory Results Using V-BLAST Space-Time Communication Architecture," *Electronic Lett.*, Vol. 35, No. 1, Jan 1999.

Problems

6.1 Use the flat earth model. Assume the following values: $|R| = .5$, $\psi = 0$, $E_0/r_1 = 1$, $h_1 = 5$ m, $h_2 = 20$ m, $d = 100$ m.

- (a) What are the arrival time delays τ_1 and τ_2 ?
- (b) Plot a family of curves for the envelope and the phase vs. $(2kh_1h_2)/d$

6.2 Use the flat earth model. Assume the following values: $|R| = .5$, $\psi = 0$, $h_1 = 5$ m, $h_2 = 20$ m, $d = 100$ m. A baseband rectangular pulse is transmitted whose width is 10 ns. The transmit pulse is shown in Fig. P6.1. The received voltage magnitude for the direct path $V_1 = 1$ V and is proportional to E_0/r_1 .

- (a) What are the arrival time delays τ_1 and τ_2 ?
- (b) Estimate the received voltage magnitude V_2 for the reflected path using spherical spreading and $|R|$.
- (c) Plot the received pulses for the direct and indirect paths showing pulse magnitudes and time delays.

6.3 Let us model an urban environment by using MATLAB to generate 100 values of the amplitude a_n , the phase α_n , and the distance r_n . Use Eqs. (6.13), (6.14), and (6.15). Generate values for the amplitudes, a_n , using `randn` and assume that $\sigma = 1$ mV. Use `rand` to generate phase values for α_n to vary between 0 and 2π . Use `rand` to define $r_n = 500 + 50 * \text{rand}(1, N)$. Find the phase ϕ_n . Plot the envelope r (mV) for $100 \text{ kHz} < f < 500 \text{ kHz}$. Let the x axis be linear and the y axis be $\log 10$.

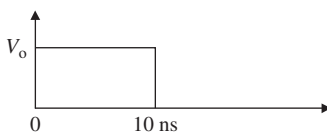


Figure P6.1 Transmit pulse.

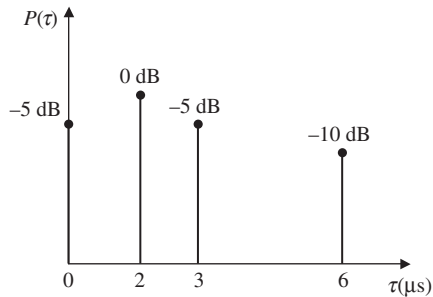


Figure P6.2 Power delay profile.

- 6.4** Calculate the outage probability for the Rayleigh process where $\sigma = 2$ mV and the threshold $p_{\text{th}} = 3 \mu\text{W}$.
- 6.5** Plot the probability density function for the Rician distribution where the Rician factor $K = -5$ dB, 5 dB, and 15 dB. Superimpose all curves on the same plot. Assume that $\sigma = 2$ mV.
- 6.6** Repeat Prob. 3 for the Rician channel where $A = 4\sigma$.
- 6.7** For the Rician-fading channel where $\sigma = 2$ mV, the direct path amplitude $A = 4$ mV, what is the probability that the received voltage envelope will exceed a threshold of 6 mV?
- 6.8** Use MATLAB to plot the envelope in Eq. (6.27) where the carrier frequency is 2 GHz, the vehicle velocity is 30 mph, the phase angles ϕ_n, θ_n are uniformly distributed, the coefficient a_n has a Gaussian distribution with zero mean and standard deviation of $\sigma = 0.002$. Let $N = 5$.
- 6.9** Calculate the multipath power gain P_T , the mean excess delay τ_0 , and the RMS delay spread σ_τ for the PDP given in Fig. P6.2.
- 6.10** Calculate the multipath power gain P_T , the mean excess delay τ_0 , and the RMS delay spread σ_τ for the PDP given in Fig. P6.3.

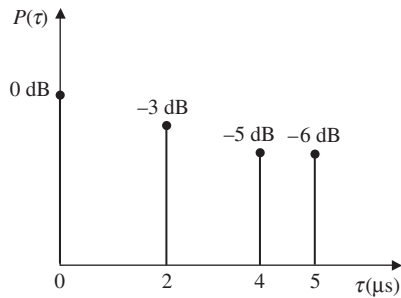


Figure P6.3 Power delay profile.

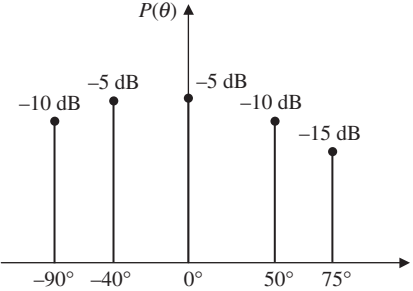


Figure P6.4 Power angular profile.

6.11 Calculate the multipath power gain P_T , the mean arrival angle θ_0 , and the RMS angular spread σ_θ for the PAP given in Fig. P6.4.

6.12 Calculate the multipath power gain P_T , the mean arrival angle θ_0 , and the RMS angular spread σ_θ for the PAP given in Fig. P6.5.

6.13 For $P(\theta)$ in Prob. 11, use the Rapport method to calculate the angular spread.

6.14 For $P(\theta)$ in Prob. 12, use the Rapport method to calculate the angular spread.

6.15 For a transmitter transmitting 1MW of power located at the rectangular coordinates (0, 250) and the receiver is located at (1km,250). Four scatterers are located at: (200, 400), (300, 150), (500, 100), (600, 400), and (800, 100). Each isotropic scatterer has a reflection coefficient of .7. Account for spherical spreading such that the power is inversely proportional to r^2 .

- (a) Derive the time delays and powers associated with each path and plot the PDP
- (b) Derive the angles of arrival and powers associated with each path and plot the PAP
- (c) Plot the two dimensional power delay angular profile given in Eq. (6.46)

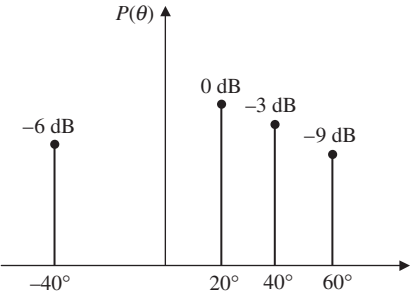


Figure P6.5 Power angular profile.

6.16 For a ring of scatterers where the circle radius $a = 20$ m and the distance between the transmit and receive antennas is 100 m. Allow $P_T = 1$ Watt

- (a) Solve for the power angular profile
- (b) Plot the PAP

6.17 For a disk of scatterers where the circle radius $a = 20$ m and the distance between the transmit and receive antennas is 100 m. Allow $P_T = 1$ Watt

- (a) Solve for the power angular profile
- (b) Plot the PAP

6.18 If the channel impulse response is given by $h_c(t) = .2\delta(t) + .4\delta(t - \tau) + .6\delta(t - 2\tau)$

- (a) Derive and plot the channel frequency response for $.5 < f\tau < 1.5$
- (b) Derive and plot the equalizer frequency response for $.5 < f\tau < 1.5$
- (c) Are these plots conjugate inverses?

6.19 If the channel impulse response is given by $h_c(t) = .6\delta(t) + .4\delta(t - 2\tau) + .2\delta(t - 4\tau)$

- (a) Derive and plot the channel frequency response for $.5 < f\tau < 1.5$
- (b) Derive and plot the equalizer frequency response for $.5 < f\tau < 1.5$
- (c) Are these plots conjugate inverses?

Chapter
7

Angle-of-Arrival Estimation

In the propagation channel topics discussed in Chap. 6, it was apparent that even for one source there are many possible propagation paths and angles of arrival. If several transmitters are operating simultaneously, each source potentially creates many multipath components at the receiver. Therefore, it is important for a receive array to be able to estimate the angles of arrival in order to decipher which emitters are present and what are their possible angular locations. This information can be used to eliminate or combine signals for greater fidelity, suppress interferers, or both.

Angle-of-arrival (AOA) estimation has also been known as *spectral* estimation, *direction of arrival* (DOA) estimation, or *bearing* estimation. Some of the earliest references refer to spectral estimation as the ability to select various frequency components out of a collection of signals. This concept was expanded to include frequency-wavenumber problems and subsequently AOA estimation. Bearing estimation is a term more commonly used in the sonar community and is AOA estimation for acoustic problems. Much of the state-of-the-art in AOA estimation has its roots in *time series analysis*, *spectrum analysis*, *periodograms*, *eigenstructure methods*, *parametric methods*, *linear prediction methods*, *beamforming*, *array processing*, and *adaptive array methods*. Some of the more useful materials include a survey paper by Godara [1], spectrum analysis by Capon [2], a review of spectral estimation by Johnson [3], an exhaustive text by Van Trees [4] and a text by Stoica and Moses [5].

7.1 Fundamentals of Matrix Algebra

Before beginning our development of AOA (spectral) estimation methods, it is important to review some matrix algebra basics. We will denote

all vectors as lower case with a bar. An example is the array vector $\bar{a}$. We will denote all matrices as upper case also with a bar such as $\bar{A}$.

7.1.1 Vector basics

Column Vector: The vector $\bar{a}$ can be denoted as a column vector or as a row vector. If $\bar{a}$ is a column vector or a single column matrix, it can be described as

$$\bar{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_M \end{bmatrix} \quad (7.1)$$

Row Vectors: If $\bar{b}$ is a row vector or a single row matrix, it can be described as

$$\bar{b} = [b_1 \quad b_2 \quad \cdots \quad b_N] \quad (7.2)$$

Vector Transpose: Any column vector can be changed into a row vector or any row vector can be changed into a column vector by the transpose operation such that

$$\bar{a}^T = [a_1 \quad a_2 \quad \cdots \quad a_M] \quad (7.3)$$

$$\bar{b}^T = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{bmatrix} \quad (7.4)$$

Vector Hermitian Transpose: The *Hermitian transpose* is the conjugate transpose of a vector denoted by the operator H . Thus the Hermitian transpose¹ of $\bar{a}$ above can be demonstrated as

$$\bar{a}^H = [a_1^* \quad a_2^* \quad \cdots \quad a_M^*] \quad (7.5)$$

$$\bar{b}^H = \begin{bmatrix} b_1^* \\ b_2^* \\ \vdots \\ b_N^* \end{bmatrix} \quad (7.6)$$

¹The Hermitian transpose is also designated by the symbol $\dagger$

Vector Dot Product (Inner Product): The *dot product* of row vector with itself is traditionally given by

$$\begin{aligned}\bar{b} \cdot \bar{b}^T &= [b_1 \quad b_2 \quad \cdots \quad b_N] \cdot \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{bmatrix} \\ &= b_1^2 + b_2^2 + \cdots + b_N^2\end{aligned}\quad (7.7)$$

Vandermonde Vector: A *Vandermonde vector* is an M -element vector such that

$$\bar{a} = \begin{bmatrix} x^0 \\ x^1 \\ \vdots \\ x^{(M-1)} \end{bmatrix}\quad (7.8)$$

Thus, the array steering vector of Eq. (4.8) is a Vandermonde vector.

7.1.2 Matrix basics

A matrix is an $M \times N$ collection of elements such that

$$\bar{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{M1} & a_{M2} & \cdots & a_{MN} \end{bmatrix}\quad (7.9)$$

where $M \times N$ is the size or *order* of the *matrix*.

Matrix Determinant The *determinant* of a square matrix can be defined by the Laplace expansion and is given by

$$\begin{aligned}|\bar{A}| &= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1M} \\ a_{21} & a_{22} & \cdots & a_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ a_{M1} & a_{M2} & \cdots & a_{MM} \end{vmatrix} \\ &= \sum_{j=1}^M a_{ij} \text{cof}(a_{ij}) \quad \text{for any row index } i \\ &= \sum_{i=1}^M a_{ij} \text{cof}(a_{ij}) \quad \text{for any column index } j\end{aligned}\quad (7.10)$$

where, $\text{cof}(a_{ij})$ is the cofactor of the element a_{ij} and is defined by $\text{cof}(a_{ij}) = (-1)^{i+j}M_{ij}$ and M_{ij} is the minor of a_{ij} . The minor is determinant of the matrix left after striking the i th row and the j th column. If any two rows or two columns of a matrix are identical, the determinant is zero. The determinant operation in MATLAB is performed by the command $\text{det}(A)$.

Example 7.1 Find the determinant of $\bar{A} = \begin{bmatrix} 1 & 2 & 0 \\ 3 & 2 & 1 \\ 5 & 1 & -1 \end{bmatrix}$

Solution Using the first row indices $|\bar{A}| = \sum_{j=1}^M a_{1j}\text{cof}(a_{1j})$. Thus

$$\begin{aligned} |\bar{A}| &= 1 \cdot (2 \cdot (-1) - 1 \cdot 1) - 2 \cdot (3 \cdot (-1) - 5 \cdot 1) + 0 \cdot (3 \cdot (-1) - 5 \cdot 1) \\ &= 13 \end{aligned}$$

The same result may be found by using the two MATLAB commands:

```
>> A = [1 2 0; 3 2 1; 5 1 -1];
>> det(A)
ans = 13
```

Matrix Addition Matrices can be *added* or *subtracted* by simply adding or subtracting the same elements of each. Thus

$$\bar{C} = \bar{A} \pm \bar{B} \Rightarrow c_{ij} = a_{ij} \pm b_{ij} \quad (7.11)$$

Matrix Multiplication *Matrix multiplication* can occur if the column index of the first equals the row index of the second. Thus an $M \times N$ matrix can be multiplied by an $N \times L$ matrix yielding an $M \times L$ matrix. The multiplication is such that

$$\bar{C} = \bar{A} \cdot \bar{B} \Rightarrow c_{ij} = \sum_{k=1}^N a_{ik}b_{kj} \quad (7.12)$$

Example 7.2 Multiply the two matrices $\bar{A} = \begin{bmatrix} 1 & -2 \\ 3 & 4 \end{bmatrix}$ $\bar{B} = \begin{bmatrix} 7 & 3 \\ -1 & 5 \end{bmatrix}$

Solution

$$\bar{A} \cdot \bar{B} = \begin{bmatrix} 9 & -7 \\ 17 & 29 \end{bmatrix}$$

This also may be accomplished in MATLAB with the following commands:

```
>> A = [1 -2; 3 4];
>> B = [7 3; -1 5];
>> A*B
```

$$\text{ans} = \begin{bmatrix} 9 & -7 \\ 17 & 29 \end{bmatrix}$$

Identity Matrix The *identity matrix*, denoted by $\bar{I}$, is defined as an $M \times M$ matrix with ones along the diagonal and zeros for all other matrix elements such that

$$\bar{I} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \quad (7.13)$$

The product of an identity matrix $\bar{I}$ with any square matrix $\bar{A}$ yields $\bar{A}$ such that $\bar{I} \cdot \bar{A} = \bar{A} \cdot \bar{I} = \bar{A}$. The identity matrix can be created in MATLAB with the command `eye(M)`. This produces an $M \times M$ identity matrix.

Cartesian Basis Vectors The columns of the identity matrix $\bar{I}$, are called *Cartesian basis vectors*. The Cartesian basis vectors are denoted by $\bar{u}_1, \bar{u}_2, \dots, \bar{u}_M$ such that $\bar{u}_1 = [1 \ 0 \ \cdots \ 0]^T$, $\bar{u}_2 = [0 \ 1 \ \cdots \ 0]^T$, $\dots$, $\bar{u}_M = [0 \ 0 \ \cdots \ 1]^T$. Thus, the identity matrix can be defined as $\bar{I} = [\bar{u}_1 \ \bar{u}_2 \ \cdots \ \bar{u}_M]$.

Trace of a Matrix The *trace* of a square matrix is the sum of the diagonal elements such that

$$\text{Tr}(\bar{A}) = \sum_{i=1}^N a_{ii} \quad (7.14)$$

The trace of a matrix in MATLAB is found by the command `trace(A)`.

Matrix Transpose The *transpose* of a matrix is the interchange of the rows and columns denoted by $\bar{A}^T$. The transpose of the product of two matrices is the product of the transposes in reverse order such that $(\bar{A} \cdot \bar{B})^T = \bar{B}^T \cdot \bar{A}^T$. The transpose in MATLAB is performed with the command `transpose(A)` or `A'`.

Matrix Hermitian Transpose The *Hermitian transpose* is the transpose of the conjugate (or the conjugate transpose) of the matrix

elements denoted by $\bar{A}^H$. The Hermitian transpose in MATLAB is performed by the operation `ctranspose(A)` or `A'`. The determinant of a matrix is the same as the determinant of its transpose thus $|\bar{A}| = |\bar{A}^T|$. The determinant of a matrix is the conjugate of its Hermitian transpose thus $|\bar{A}| = |\bar{A}^H|^*$. The Hermitian transpose of the product of two matrices is the product of the Hermitian transposes in reverse order such that $(\bar{A} \cdot \bar{B})^H = \bar{B}^H \cdot \bar{A}^H$. This is an important property which will be used later in the text.

Inverse of a Matrix The inverse of a matrix is defined such that $\bar{A} \cdot \bar{A}^{-1} = \bar{I}$ where $\bar{A}^{-1}$ is the inverse of $\bar{A}$. The matrix $\bar{A}$ has an inverse provided that $|\bar{A}| \neq 0$. We define the cofactor matrix such that $\bar{C} = \text{cof}(\bar{A}) = [(-1)^{i+j} |\bar{A}_{ij}|]$ and $\bar{A}_{ij}$ is the remaining matrix after striking row i and column j . The inverse of a matrix in MATLAB is given by `inv(A)`. Mathematically, the matrix inverse is given by

$$\bar{A}^{-1} = \frac{\bar{C}^T}{|\bar{A}|} \quad (7.15)$$

Example 7.3 Find the inverse of the matrix $\bar{A} = \begin{bmatrix} 1 & 3 \\ -2 & 5 \end{bmatrix}$

Solution First find the cofactor matrix $\bar{C}$ and the determinant of $\bar{A}$.

$$\bar{C} = \begin{bmatrix} 5 & 2 \\ -3 & 1 \end{bmatrix} \quad |\bar{A}| = 11$$

The inverse of $\bar{A}$ is then given as

$$\bar{A}^{-1} = \frac{\begin{bmatrix} 5 & 2 \\ -3 & 1 \end{bmatrix}^T}{11} = \begin{bmatrix} .4545 & -.2727 \\ .1818 & .0909 \end{bmatrix}$$

This problem can also be easily solved in MATLAB using the following commands:

```
>> A = [1 3;-2 5];
>> inv(A)
```

$$\text{ans} = \begin{bmatrix} .4545 & -.2727 \\ .1818 & .0909 \end{bmatrix}$$

Eigenvalues and eigenvectors of a matrix. The German word *eigen* means *appropriate* or *peculiar*. Thus, the *eigenvalues* and *eigenvectors* of a matrix are the appropriate or peculiar values that satisfy a homogeneous condition. The values of λ , that are eigenvalues of the $N \times N$ square matrix $\bar{A}$, must satisfy the following condition

$$|\lambda \bar{I} - \bar{A}| = 0 \quad (7.16)$$

The determinant above is called the *characteristic determinant* producing a polynomial of order N in λ having N roots. In other words

$$|\lambda \bar{I} - \bar{A}| = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_N) \quad (7.17)$$

Each of the eigenvalues $(\lambda_1, \lambda_2, \dots, \lambda_N)$ satisfies Eq. (7.16). We may now also define eigenvectors associated with the matrix $\bar{A}$. If the $N \times N$ matrix $\bar{A}$ has N unique eigenvalues λ_j , it has N eigenvectors $\bar{e}_j$ satisfying the following homogeneous equation:

$$(\lambda_j \bar{I} - \bar{A})\bar{e}_j = 0 \quad (7.18)$$

In MATLAB, one can find the eigenvectors and eigenvalues of a matrix A by the command $[EV, V] = \text{eig}(A)$. The eigenvectors are the columns of the matrix EV and the corresponding eigenvalues are the diagonal elements of the matrix V . The MATLAB command $\text{diag}(V)$ creates a vector of the eigenvalues along the diagonal of V .

Example 7.4 Use MATLAB to find the eigenvectors and eigenvalues of the matrix $\begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}$.

Solution The following lines create the matrix and compute the eigenvectors and eigenvalues:

```
>> A=[1 2;3 5];
>> [EV,V]=eig(A);
>> EV
```

```
EV =
-0.8646    -0.3613
0.5025    -0.9325
```

```
>> diag(V)
```

```
ans =
-0.1623
6.1623
```

The first column of EV is the first eigenvector and the corresponding eigenvalue is $\lambda_1 = -0.1623$. The second column of EV is the second eigenvector and the corresponding eigenvalue is $\lambda_2 = 6.1623$.

These simple vector and matrix procedures will assist us in using MATLAB to solve for AOA estimation algorithms that will be described in Sec. 7.3.

7.2 Array Correlation Matrix

Many of the AOA algorithms rely on the array correlation matrix. In order to understand the array correlation matrix, let us begin with a description of the array, the received signal, and the additive noise.

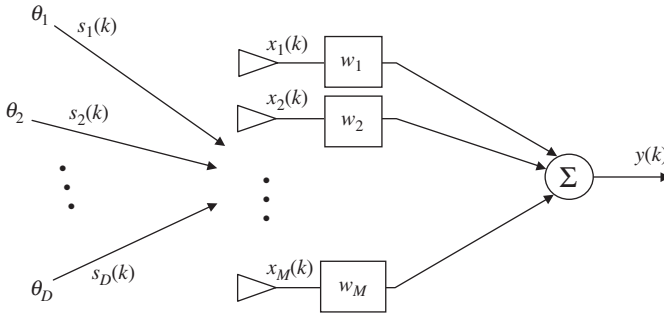


Figure 7.1 M -element array with arriving signals.

Figure 7.1 depicts a receive array with incident planewaves from various directions.

Figure 7.1 shows D signals arriving from D directions. They are received by an array of M elements with M potential weights. Each received signal $x_m(k)$ includes additive, zero mean, Gaussian noise. Time is represented by the k th time sample. Thus, the array output y can be given in the following form:

$$y(k) = \bar{w}^T \cdot \bar{x}(k) \quad (7.19)$$

where

$$\begin{aligned} \bar{x}(k) &= [\bar{a}(\theta_1) \quad \bar{a}(\theta_2) \cdots \bar{a}(\theta_D)] \cdot \begin{bmatrix} s_1(k) \\ s_2(k) \\ \vdots \\ s_D(k) \end{bmatrix} + \bar{n}(k) \\ &= \bar{A} \cdot \bar{s}(k) + \bar{n}(k) \end{aligned} \quad (7.20)$$

and

$$\bar{w} = [w_1 \quad w_2 \quad \cdots \quad w_M]^T = \text{array weights}$$

$\bar{s}(k)$ = vector of incident complex monochromatic signals at time k

$\bar{n}(k)$ = noise vector at each array element m , zero mean, variance σ_n^2

$\bar{a}(\theta_i)$ = M -element array steering vector for the θ_i direction of arrival

$\bar{A} = [\bar{a}(\theta_1) \bar{a}(\theta_2) \cdots \bar{a}(\theta_D)]$ $M \times D$ matrix of steering vectors $\bar{a}\theta_i$

Thus, each of the D -complex signals arrives at angles θ_i and is intercepted by the M antenna elements. It is initially assumed that the arriving signals are monochromatic and the number of arriving signals $D < M$. It is understood that the arriving signals are time varying and thus our calculations are based upon time snapshots of the

incoming signal. Obviously if the transmitters are moving, the matrix of steering vectors is changing with time and the corresponding arrival angles are changing. Unless otherwise stated, the time dependence will be suppressed in Eqs. (7.19) and (7.20). In order to simplify the notation let us define the $M \times M$ array correlation matrix $\bar{R}_{xx}$ as

$$\begin{aligned}\bar{R}_{xx} &= E[\bar{x} \cdot \bar{x}^H] = E[(\bar{A}\bar{s} + \bar{n})(\bar{s}^H \bar{A}^H + \bar{n}^H)] \\ &= \bar{A}E[\bar{s} \cdot \bar{s}^H]\bar{A}^H + E[\bar{n} \cdot \bar{n}^H] \\ &= \bar{A}\bar{R}_{ss}\bar{A}^H + \bar{R}_{nn}\end{aligned}\quad (7.21)$$

where $\bar{R}_{ss} = D \times D$ source correlation matrix

$\bar{R}_{nn} = \sigma_n^2 \bar{I} = M \times M$ noise correlation matrix

$\bar{I} = N \times N$ identity matrix

The array correlation matrix $\bar{R}_{xx}$ and the source correlation matrix $\bar{R}_{ss}$ are found by the expected value of the respective absolute values squared (i.e., $\bar{R}_{xx} = E[\bar{x} \cdot \bar{x}^H]$ and $\bar{R}_{ss} = E[\bar{s} \cdot \bar{s}^H]$). If we do not know the exact statistics for the noise and signals, but we can assume that the process is ergodic, then we can approximate the correlation by use of a time-averaged correlation. In that case the correlation matrices are defined by

$$\hat{R}_{xx} \approx \frac{1}{K} \sum_{k=1}^K \bar{x}(k) \bar{x}^H(k) \quad \hat{R}_{ss} \approx \frac{1}{K} \sum_{k=1}^K \bar{s}(k) \bar{s}^H(k) \quad \hat{R}_{nn} \approx \frac{1}{K} \sum_{k=1}^K \bar{n}(k) \bar{n}^H(k)$$

When the signals are uncorrelated, $\bar{R}_{ss}$, obviously has to be a diagonal matrix because off-diagonal elements have no correlation. When the signals are partly correlated, $\bar{R}_{ss}$ is nonsingular. When the signals are coherent, $\bar{R}_{ss}$ becomes singular because the rows are linear combinations of each other [5]. The matrix of steering vectors, $\bar{A}$, is an $M \times D$ matrix where all columns are different. Their structure is Vandermonde and hence the columns are independent [6, 7]. Often in the literature, the array correlation matrix is referred to as the covariance matrix. This is only true if the mean values of the signals and noise are zero. In that case, the covariance and the correlation matrices are identical. The arriving signal mean value must necessarily be zero because antennas cannot receive d.c. signals. The noise inherent in the receiver may or may not have zero mean depending on the source of the receiver noise.

There is much useful information to be discovered in the eigenanalysis of the array correlation matrix. Details of the eigenstructure are described in Godara [1] and are repeated here. Given M -array elements with D -narrowband signal sources and uncorrelated noise we can make some assumptions about the properties of the correlation matrix. First, $\bar{R}_{xx}$ is an $M \times M$ Hermitian matrix. A Hermitian matrix is equal to its

complex conjugate transpose such that $\bar{R}_{xx} = \bar{R}_{xx}^H$. The array correlation matrix has M eigenvalues $(\lambda_1, \lambda_2, \dots, \lambda_M)$ along with M associated eigenvectors $\bar{E} = [\bar{e}_1 \bar{e}_2 \dots \bar{e}_M]$. If the eigenvalues are sorted from smallest to largest, we can divide the matrix $\bar{E}$ into two subspaces such that $\bar{E} = [\bar{E}_N \bar{E}_S]$. The first subspace $\bar{E}_N$ is called the *noise subspace* and is composed of $M - D$ eigenvectors associated with the noise. For uncorrelated noise, the eigenvalues are given as $\lambda_1 = \lambda_2 = \dots = \lambda_{M-D} = \sigma_n^2$. The second subspace $\bar{E}_S$ is called the *signal subspace* and is composed of D eigenvectors associated with the arriving signals. The noise subspace is an $M \times (M - D)$ matrix. The signal subspace is an $M \times D$ matrix.

The goal of AOA estimation techniques is to define a function that gives an indication of the angles of arrival based upon maxima vs. angle. This function is traditionally called the *pseudospectrum* $P(\theta)$ and the units can be in energy or in watts (or at times energy or watts squared). There are several potential approaches to defining the pseudospectrum via beamforming, the array correlation matrix, eigenanalysis, linear prediction, minimum variance, maximum likelihood, min-norm, MUSIC, root-MUSIC, and many more approaches that will not be addressed in this chapter. Both Stoica and Moses [5] and Van Trees [4] give an in-depth explanation of many of these possible approaches. We will summarize some of the more popular pseudospectra solutions in the next section.

7.3 AOA Estimation Methods

7.3.1 Bartlett AOA estimate

If the array is uniformly weighted, we can define the Bartlett AOA estimate [8] as

$$P_B(\theta) = \bar{a}^H(\theta) \bar{R}_{xx} \bar{a}(\theta) \quad (7.22)$$

The Bartlett AOA estimate is the spatial version of an averaged periodogram and is a beamforming AOA estimate. Under the conditions where $\bar{s}$ represents uncorrelated monochromatic signals and there is no system noise, Eq. (7.22) is equivalent to the following long-hand expression:

$$P_B(\theta) = \left| \sum_{i=1}^D \sum_{m=1}^M e^{j(m-1)kd(\sin \theta - \sin \theta_i)} \right|^2 \quad (7.23)$$

The periodogram is thus equivalent to the spatial finite Fourier transform of all arriving signals. This is also equivalent to adding all beam-steered array factors for each angle of arrival and finding the absolute value squared.

Example 7.5 Use MATLAB to plot the pseudo-spectrum using the Bartlett estimate for an $M = 6$ element array. With element spacing $d = \lambda/2$, uncorrelated, equal amplitude sources, (s_1, s_2) , and $\sigma_n^2 = .1$, and the two different pairs of arrival angles given by $\pm 10^\circ$ and $\pm 5^\circ$, assume ergodicity.

Solution From the information given we can find the following:

$$\bar{s} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \bar{a}(\theta) = [1 \quad e^{j\pi \sin \theta} \quad \dots \quad e^{j5\pi \sin \theta}]^T$$

$$\bar{A} = [\bar{a}(\theta_1) \quad \bar{a}(\theta_2)] \quad \bar{R}_{ss} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Applying Eq. (7.21) we can find $\bar{R}_{xx}$ for both sets of angles. Substituting $\bar{R}_{xx}$ into Eq. (7.22) and using MATLAB, we can plot the pseudospectrum as shown in Figs. 7.2a and b.

Recalling the half-power beamwidth of a linear array from Chap. 4, Eq. (4.21), we can estimate the beamwidth of this $M = 6$ element array to be $\approx 8.5^\circ$. Thus, the two sources, which are 20° apart are resolvable with the Bartlett approach. The two sources, which are 10° apart are not resolvable. Herein lies one of the limitations of the Bartlett approach to AOA estimation: the ability to resolve angles is limited by the array half-power beamwidth. An increase in resolution requires a larger array. For large array lengths with $d = \lambda/2$ spacing, the AOA resolution is approximately $1/M$. Thus, $1/M$ is the AOA resolution limit of a periodogram and in the case above is an indicator of the resolution of the Bartlett method. It should be noted that when two emitters are separated by an angle wider than the array resolution, they can be resolved but a bias is introduced. This bias cause the peaks to deviate from the true AOA. This bias asymptotically decreases as the array length increases.

7.3.2 Capon AOA estimate

The Capon AOA estimate [2, 4] is known as a *minimum variance distortionless response* (MVDR). It is also alternatively a maximum likelihood estimate of the power arriving from one direction while all other sources are considered as interference. Thus the goal is to maximize the signal-to-interference ratio (SIR) while passing the signal of interest undistorted in phase and amplitude. The source correlation matrix $\bar{R}_{ss}$ is assumed to be diagonal. This maximized SIR is accomplished with a set of array weights ($\bar{w} = [w_1 w_2 \dots w_M]^T$) as shown in Fig. 7.1, where the array weights are given by

$$\bar{w} = \frac{\bar{R}_{xx}^{-1} \bar{a}(\theta)}{\bar{a}^H(\theta) \bar{R}_{xx}^{-1} \bar{a}(\theta)} \quad (7.24)$$

where $\bar{R}_{xx}$ is the unweighted array correlation matrix.

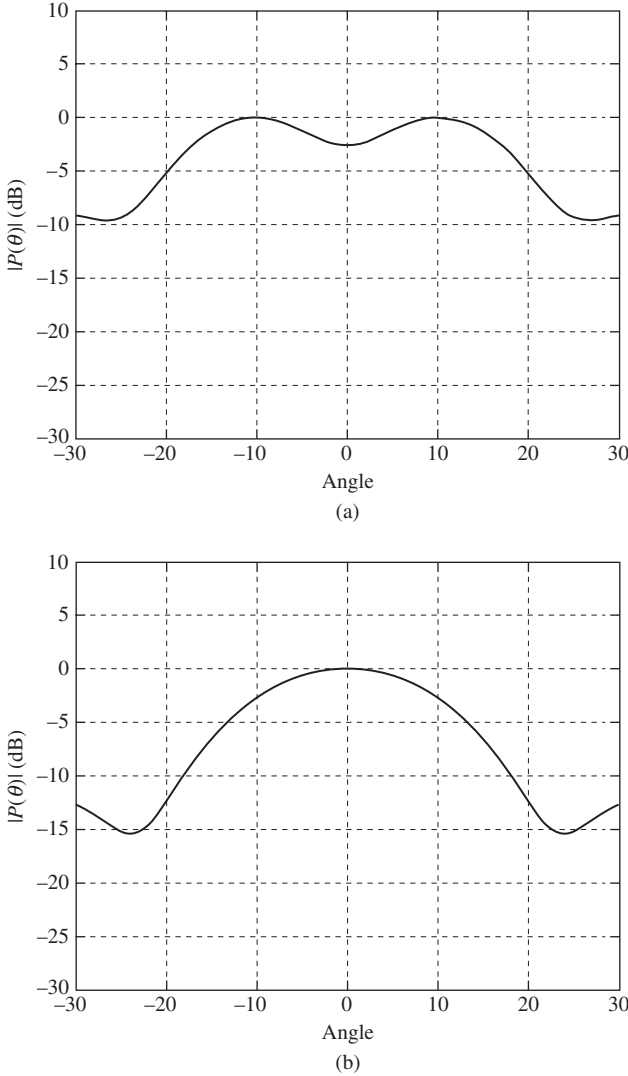


Figure 7.2 (a) Bartlett pseudospectrum for $\theta_1 = -10^\circ$, $\theta_2 = 10^\circ$.
(b) Bartlett pseudospectrum for $\theta_1 = -5^\circ$, $\theta_2 = 5^\circ$.

Substituting the weights of Eq. (7.24) into the array of Fig. 7.1, we can then find that the pseudospectrum is given by

$$P_C(\theta) = \frac{1}{\bar{\mathbf{a}}^H(\theta) \bar{\mathbf{R}}_{xx}^{-1} \bar{\mathbf{a}}(\theta)} \quad (7.25)$$

Example 7.6 Use MATLAB to plot the pseudo-spectrum using the Capon estimate for an $M = 6$ element array. With element spacing $d = \lambda/2$, uncorrelated,

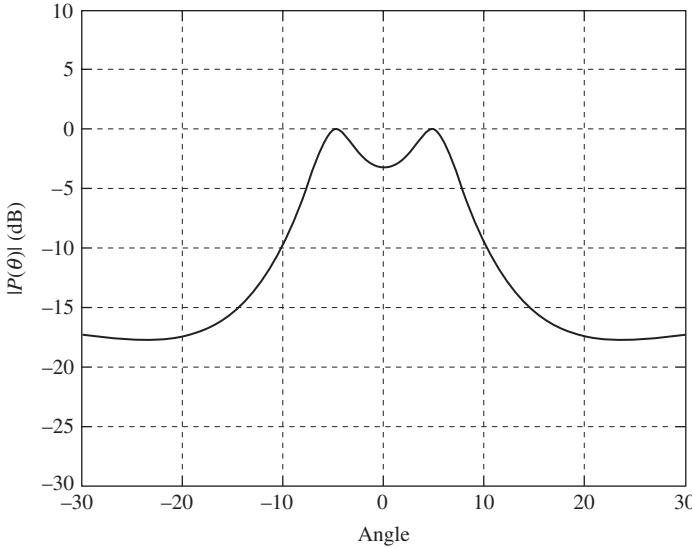


Figure 7.3 Capon (ML) pseudospectrum for $\theta_1 = -5^\circ$, $\theta_2 = 5^\circ$.

equal amplitude sources, (s_1, s_2) , and $\sigma_n^2 = .1$, and the pair of arrival angles given by $\pm 5^\circ$, assume ergodicity.

Solution We can use the same array correlation matrix as was found in Example 7.5. Using MATLAB we produce Fig. 7.3.

It is clear that the Capon AOA estimate has much greater resolution than the Bartlett AOA estimate. In the case where the competing sources are highly correlated, the Capon resolution can actually become worse. The derivation of the Capon (ML) weights was conditioned upon considering that all other sources are interferers. If the multiple signals can be considered as multipath signals, with Rayleigh amplitude and uniform phase, then the uncorrelated condition is met and the Capon estimate will work.

The advantage of the Bartlett and Capon estimation methods is that these are nonparametric solutions and one does not need an a priori knowledge of the specific statistical properties.

7.3.3 Linear prediction AOA estimate

The goal of the linear prediction method is to minimize the prediction error between the output of the m th sensor and the actual output [3, 9]. Our goal is to find the weights that minimize the mean-squared prediction error. In a similar vein as Eq. (7.24), the solution for the array weights is given as

$$\bar{w}_m = \frac{\bar{R}_{xx}^{-1} \bar{u}_m}{\bar{u}_m^T \bar{R}_{xx}^{-1} \bar{u}_m} \quad (7.26)$$

where $\bar{u}_m$ is the *Cartesian basis vector* which is the m th column of the $M \times M$ identity matrix.

Upon substitution of these array weights into the calculation of the pseudo-spectrum, it can be shown that

$$P_{LP_m}(\theta) = \frac{\bar{u}_m^T \bar{R}_{xx}^{-1} \bar{u}_m}{|\bar{u}_m^T \bar{R}_{xx}^{-1} \bar{a}(\theta)|^2} \quad (7.27)$$

The particular choice for which m th element output for prediction is random. Although the choice made can dramatically affect the final resolution. If the array center element is chosen, the linear combination of the remaining sensor elements might provide a better estimate because the other array elements are spaced about the phase center of the array [3]. This would suggest that odd array lengths might provide better results than even arrays because the center element is precisely at the array phase center.

This linear prediction technique is sometimes referred to as an *autoregressive* method [4]. It has been argued that the spectral peaks using linear prediction are proportional to the square of the signal power [3]. This is true in Example 7.7.

Example 7.7 Use MATLAB to plot the pseudo-spectrum using the linear predictive estimate for an $M = 6$ element array. With element spacing $d = \lambda/2$, uncorrelated, equal amplitude sources, (s_1, s_2) , and $\sigma_n^2 = .1$, and the pair of arrival angles given by $\pm 5^\circ$, choose the 3rd element of the array as the reference element such that the Cartesian basis vector is $\bar{u}_3 = [0 \ 0 \ 1 \ 0 \ 0 \ 0]^T$. Assume ergodicity.

Solution The pseudospectra is given as $P_{LP_3}(\theta) = \frac{\bar{u}_3^T \bar{R}_{xx}^{-1} \bar{u}_3}{|\bar{u}_3^T \bar{R}_{xx}^{-1} \bar{a}(\theta)|^2}$ and is plotted in Fig. 7.4.

It is very obvious that under these conditions, the linear predictive method provides superior performance over both the Bartlett estimate and the Capon estimate. The efficacy of the performance is dependent on the array element chosen and the subsequent $\bar{u}_n$ vector. When one selects the arrival signals to have different amplitudes, the linear predictive spectral peaks reflect the relative strengths of the incoming signals. Thus, the linear predictive method not only provides AOA information but it also provides signal strength information.

7.3.4 Maximum entropy AOA estimate

The maximum entropy method is attributed to Burg [10, 11]. A further explanation of the maximum entropy approach is given in [1, 12]. The goal is to find a pseudospectrum that maximizes the entropy function subject to constraints. The details of the Burg derivation can be found

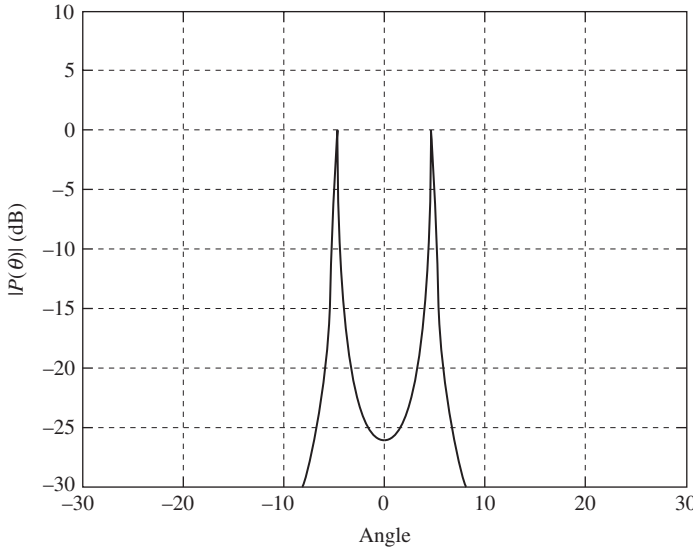


Figure 7.4 Linear predictive pseudospectrum for $\theta_1 = -5^\circ$, $\theta_2 = 5^\circ$.

in the references discussed previously. The pseudospectrum is given by

$$P_{ME_j}(\theta) = \frac{1}{\tilde{\mathbf{a}}(\theta)^H \tilde{\mathbf{c}}_j \tilde{\mathbf{c}}_j^H \tilde{\mathbf{a}}(\theta)} \quad (7.28)$$

where $\tilde{\mathbf{c}}_j$ is the j th column of the inverse array correlation matrix ($\bar{\mathbf{R}}_{xx}^{-1}$).

Example 7.8 Use MATLAB to plot the pseudospectrum using the maximum entropy AOA estimate for an $M = 6$ -element array, element spacing $d = \lambda/2$, uncorrelated, equal amplitude sources, (s_1, s_2) , and $\sigma_n^2 = .1$, and the pair of arrival angles given by $\pm 5^\circ$. Choose the 3rd column ($\tilde{\mathbf{c}}_3$) of the array correlation matrix to satisfy Eq. (7.28). Assume ergodicity.

Solution The pseudospectrum is given as plotted in Fig. 7.5

It should be noted that the maximum entropy method, when we select the $\tilde{\mathbf{c}}_3$ column from $\bar{\mathbf{R}}_{xx}^{-1}$, gives the same pseudospectra as the linear predictive method. The choice of $\tilde{\mathbf{c}}_j$ can dramatically effect the resolution achieved. The center columns of the inverse array correlation matrix tend to give better results under the conditions assumed in this chapter.

7.3.5 Pisarenko harmonic decomposition AOA estimate

The *Pisarenko harmonic decomposition* (PHD) AOA estimate is named after the Russian mathematician who devised this minimum mean-squared error approach [13, 14]. The goal is to minimize the

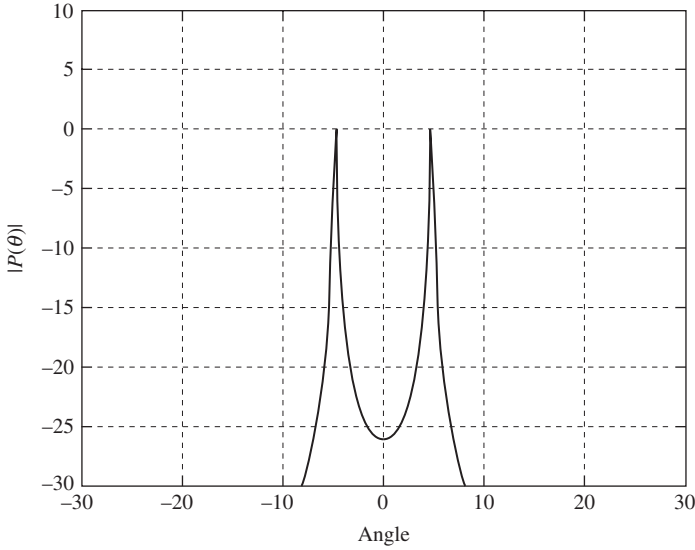


Figure 7.5 Maximum entropy pseudospectrum for $\theta_1 = -5^\circ$, $\theta_2 = 5^\circ$.

mean-squared error of the array output under the constraint that the norm of the weight vector be equal to unity. The eigenvector that minimizes the mean-squared error corresponds to the smallest eigenvalue. For an $M = 6$ -element array, with two arriving signals, there will be two eigenvectors associated with the signal and four eigenvectors associated with the noise. The corresponding PHD pseudospectrum is given by

$$P_{PHD}(\theta) = \frac{1}{|\bar{a}^H(\theta)\bar{e}_1|^2} \tag{7.29}$$

where $\bar{e}_1$ is the eigenvector associated with the smallest eigenvalue λ_1 .

Example 7.9 Use MATLAB to plot the pseudo-spectrum using the Pisarenko Harmonic Decomposition estimate for an $M = 6$ element array, element spacing $d = \lambda/2$, uncorrelated, equal amplitude sources, (s_1, s_2) , and $\sigma_n^2 = .1$, and the pair of arrival angles given by $\pm 5^\circ$. Choose the first noise eigenvector to produce the pseudospectrum

Solution After finding the array correlation matrix, we can use the `eig()` command in MATLAB to find the eigenvectors and corresponding eigenvalues. The eigenvalues are given by $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = \sigma_n^2 = .1$, $\lambda_5 = 2.95$, $\lambda_6 = 9.25$.

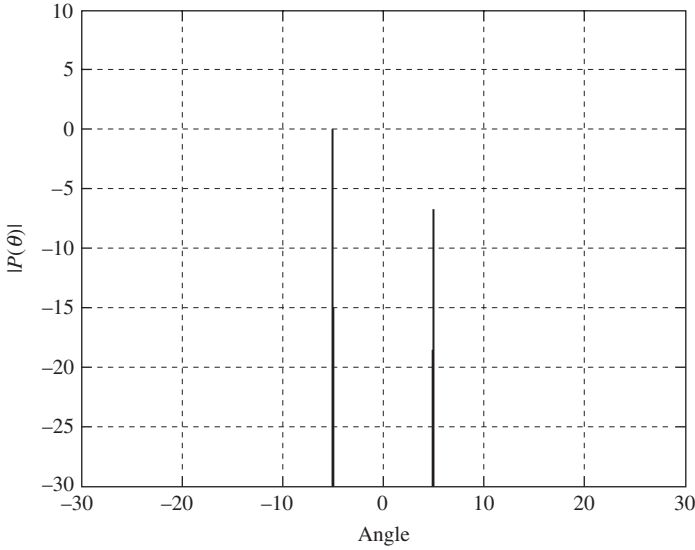


Figure 7.6 Pisarenko Harmonic Decomposition pseudospectrum for $\theta_1 = -5^\circ$, $\theta_2 = 5^\circ$.

The first eigenvector associated with λ_1 is given by

$$\bar{e}_1 = \begin{bmatrix} -0.143 \\ -0.195 \\ 0.065 \\ 0.198 \\ 0.612 \\ -0.723 \end{bmatrix}$$

Substituting this eigenvector in Eq. (7.29) we can plot Fig. 7.6

The Pisarenko peaks are not an indication of the signal amplitudes. These peaks are the roots of the polynomial in the denominator of Eq. (7.29). It is clear that for this example, the Pisarenko solution has the best resolution.

7.3.6 Min-norm AOA estimate

The minimum-norm method was developed by Reddi [15] and Kumaresan and Tufts [16]. This method is also lucidly explained by Ermolaev and Gershman [17]. The min-norm method is only relevant for *uniform linear arrays* (ULA). The min-norm algorithm optimizes the weight vector by solving the optimization problem where

$$\min_{\bar{w}} \bar{w}^H \bar{w} \quad \bar{E}_S^H \bar{w} = 0 \quad \bar{w}^H \bar{u}_1 = 1 \quad (7.30)$$

where $\bar{w}$ = array weights

$\bar{E}_S$ = subspace of D signal eigenvectors = $[\bar{e}_{M-D+1} \ \bar{e}_{M-D+2} \cdots \bar{e}_M]$

M = number of array elements

D = number of arriving signals

$\bar{u}_1$ = Cartesian basis vector (first column of the $M \times M$ identity matrix)
 $= [1 \ 0 \ 0 \cdots 0]^T$

The solution to the optimization yields the min-norm pseudospectrum

$$P_{MN}(\theta) = \frac{(\bar{u}_1^T \bar{E}_N \bar{E}_N^H \bar{u}_1)^2}{|\bar{a}(\theta)^H \bar{E}_N \bar{E}_N^H \bar{u}_1|^2} \quad (7.31)$$

where $\bar{E}_N$ = subspace of $M - D$ noise eigenvectors = $[\bar{e}_1 \ \bar{e}_2 \cdots \bar{e}_{M-D}]$

$\bar{a}(\theta)$ = array steering vector

Since the numerator term in Eq. (7.31) is a constant, we can normalize the pseudospectrum such that

$$P_{MN}(\theta) = \frac{1}{|\bar{a}(\theta)^H \bar{E}_N \bar{E}_N^H \bar{u}_1|^2} \quad (7.32)$$

Example 7.10 Use MATLAB to plot the pseudo-spectrum using the min-norm AOA estimate for an $M = 6$ element array, with element spacing $d = \lambda/2$, uncorrelated, equal amplitude sources, (s_1, s_2) , and $\sigma_n^2 = .1$, and the pair of arrival angles given by $\pm 5^\circ$. Use all noise eigenvectors to construct the noise subspace $\bar{E}_N$.

Solution After finding the array correlation matrix, we can use the `eig()` command in MATLAB to find the eigenvectors and corresponding eigenvalues. The eigenvalues are broken up into two groups. There are eigenvectors associated with the noise eigenvalues given by $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = \sigma_n^2 = .1$. There are eigenvectors associated with the signal eigenvalues $\lambda_5 = 2.95$ and $\lambda_6 = 9.25$. The subspace created by the $M - D = 4$ noise eigenvectors is given as

$$\bar{E}_N = \begin{bmatrix} -0.14 & -0.56 & -0.21 & 0.27 \\ -0.2 & 0.23 & 0.22 & -0.75 \\ 0.065 & 0.43 & 0.49 & 0.58 \\ 0.2 & 0.35 & -0.78 & 0.035 \\ 0.61 & -0.51 & 0.22 & -0.15 \\ -0.72 & -0.25 & 0 & 0.083 \end{bmatrix}$$

Applying this information to Eq. (7.32), we can plot the angular spectra in Fig. 7.7

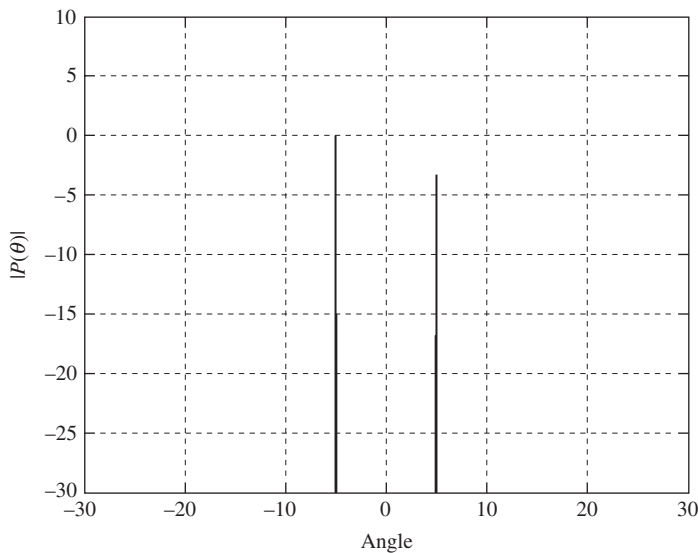


Figure 7.7 Min-norm pseudospectrum for $\theta_1 = -5^\circ, \theta_2 = 5^\circ$.

It should be noted that the pseudospectrum from the min-norm method is almost identical to the PHD pseudospectrum. The min-norm method combines all noise eigenvectors whereas the PHD method only uses the first noise eigenvector.

7.3.7 MUSIC AOA estimate

MUSIC is an acronym which stands for Multiple Signal Classification. This approach was first posed by Schmidt [18] and is a popular high resolution eigenstructure method. MUSIC promises to provide unbiased estimates of the number of signals, the angles of arrival, and the strengths of the waveforms. MUSIC makes the assumption that the noise in each channel is uncorrelated making the noise correlation matrix diagonal. The incident signals may be somewhat correlated creating a nondiagonal signal correlation matrix. However, under high signal correlation the traditional MUSIC algorithm breaks down and other methods must be implemented to correct this weakness. These methods will be discussed later in this chapter.

One must know in advance the number of incoming signals or one must search the eigenvalues to determine the number of incoming signals. If the number of signals is D , the number of signal eigenvalues and eigenvectors is D , and the number of noise eigenvalues and eigenvectors is $M - D$ (M is the number of array elements). Because MUSIC

exploits the noise eigenvector subspace, it is sometimes referred to as a *subspace method*.

As before we calculate the array correlation matrix assuming uncorrelated noise with equal variances.

$$\bar{R}_{xx} = \bar{A}\bar{R}_{ss}\bar{A}^H + \sigma_n^2 \bar{I} \quad (7.33)$$

We next find the eigenvalues and eigenvectors for $\bar{R}_{xx}$. We then produce D eigenvectors associated with the signals and $M - D$ eigenvectors associated with the noise. We choose the eigenvectors associated with the smallest eigenvalues. For uncorrelated signals, the smallest eigenvalues are equal to the variance of the noise. We can then construct the $M \times (M - D)$ dimensional subspace spanned by the noise eigenvectors such that

$$\bar{E}_N = [\bar{e}_1 \quad \bar{e}_2 \quad \cdots \quad \bar{e}_{M-D}] \quad (7.34)$$

The noise subspace eigenvectors are orthogonal to the array steering vectors at the angles of arrival $\theta_1, \theta_2, \dots, \theta_D$. Because of this orthogonality condition, one can show that the Euclidean distance $d^2 = \bar{a}(\theta)^H \bar{E}_N \bar{E}_N^H \bar{a}(\theta) = 0$ for each and every arrival angle $\theta_1, \theta_2, \dots, \theta_D$. Placing this distance expression in the denominator creates sharp peaks at the angles of arrival. The MUSIC pseudospectrum is now given as

$$P_{MU}(\theta) = \frac{1}{|\bar{a}(\theta)^H \bar{E}_N \bar{E}_N^H \bar{a}(\theta)|} \quad (7.35)$$

Example 7.11 Use MATLAB to plot the pseudospectrum using the MUSIC AOA estimate for an $M = 6$ element array. With element spacing $d = \lambda/2$, uncorrelated, equal amplitude sources, (s_1, s_2) , and $\sigma_n^2 = .1$, and the pair of arrival angles given by $\pm 5^\circ$. Use all noise eigenvectors to construct the noise subspace $\bar{E}_N$.

Solution After finding the array correlation matrix, we can use the `eig()` command in MATLAB to find the eigenvectors and corresponding eigenvalues. The eigenvalues are given by $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = \sigma_n^2 = .1$, $\lambda_5 = 2.95$, and $\lambda_6 = 9.25$. The eigenvalues and eigenvectors can be sorted in MATLAB from the least to the greatest by the following commands:

```
[V,Dia] = eig(Rxx);

[Y,Index] = sort(diag(Dia));

EN = V(:,Index(1:M-D));
```

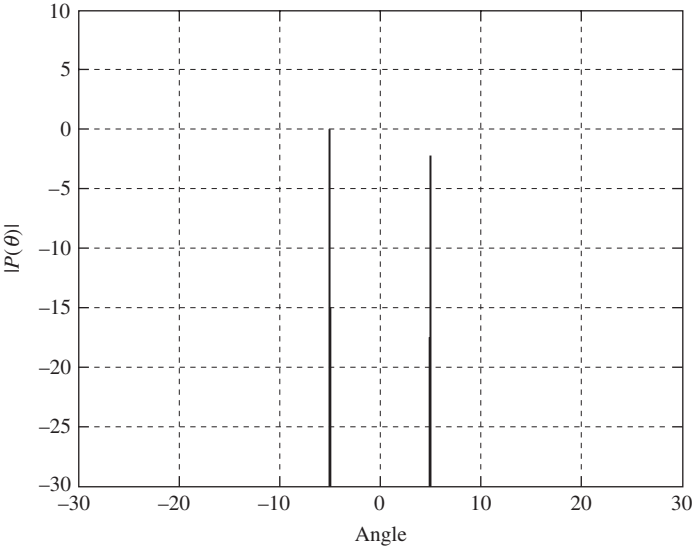


Figure 7.8 MUSIC pseudospectrum for $\theta_1 = -5^\circ$, $\theta_2 = 5^\circ$.

The subspace created by the $M - D = 4$ noise eigenvectors again is given as

$$\bar{E}_N = \begin{bmatrix} -0.14 & -0.56 & -0.21 & 0.27 \\ -0.2 & 0.23 & 0.22 & -0.75 \\ 0.065 & 0.43 & 0.49 & 0.58 \\ 0.2 & 0.35 & -0.78 & 0.035 \\ 0.61 & -0.51 & 0.22 & -0.15 \\ -0.72 & -0.25 & 0 & 0.083 \end{bmatrix}$$

Applying this information to Eq. (7.35), we can plot the angular spectra in Fig. 7.8

Under the conditions stated for the Pisarenko harmonic decomposition, the min-norm method, and the MUSIC method, the solutions all have similar resolution. It should be understood that in all examples discussed earlier, it was assumed that the array correlation matrix was of the form given in Eq. (7.33), that the noise variance for all elements was identical, and that the different signals were completely uncorrelated. In the case where the source correlation matrix is not diagonal, or the noise variances vary, the plots can change dramatically and the resolution will diminish.

In the more practical application we must collect several time samples of the received signal plus noise, assume ergodicity, and estimate the correlation matrices via time averaging. We can repeat Eq. (7.33) without assuming that

we know the signal statistics

$$\begin{aligned}\hat{R}_{xx} &= E[\bar{x}(k) \cdot \bar{x}^H(k)] \approx \frac{1}{K} \sum_{k=1}^K \bar{x}(k) \cdot \bar{x}^H(k) \\ &\approx \bar{A} \hat{R}_{ss} \bar{A}^H + \bar{A} \hat{R}_{sn} + \hat{R}_{ns} \bar{A}^H + \hat{R}_{nn}\end{aligned}\quad (7.36)$$

where

$$\begin{aligned}\hat{R}_{ss} &= \frac{1}{K} \sum_{k=1}^K \bar{s}(k) \bar{s}^H(k) & \hat{R}_{sn} &= \frac{1}{K} \sum_{k=1}^K \bar{s}(k) \bar{n}^H(k) \\ \hat{R}_{ns} &= \frac{1}{K} \sum_{k=1}^K \bar{n}(k) \bar{s}^H(k) & \hat{R}_{nn} &= \frac{1}{K} \sum_{k=1}^K \bar{n}(k) \bar{n}^H(k)\end{aligned}$$

Example 7.12 Use MATLAB to plot the pseudospectrum using the MUSIC AOA estimate for an $M = 6$ element array. With element spacing $d = \lambda/2$, the pair of arrival angles are given by $\pm 5^\circ$. Assume binary *Walsh-like* signals of amplitude 1, but with only K finite signal samples. Assume Gaussian distributed noise of $\sigma_n^2 = .1$ but with only K finite noise samples. Also assume the process is ergodic and collect $K = 100$ time samples ($k = 1, 2, \dots, K$) of the signal such that $s = \text{sign}(\text{randn}(M, K))$ and the noise such that $n = \text{sqrt}(\text{sig2}) * \text{randn}(M, K)$ ($\text{sig2} = \sigma_n^2$). Calculate all correlation matrices via time averaging as defined in Eq. (7.36). This can be accomplished in MATLAB by the commands $R_{ss} = s * s' / K$, $R_{ns} = n * s' / K$, $R_{sn} = s * n' / K$, and $R_{nn} = n * n' / K$. Assume a pair of arrival angles given by $\pm 5^\circ$. Use all noise eigenvectors to construct the noise subspace $\bar{E}_N$ and find the pseudospectrum. (*Important:* MATLAB will not order the eigenvalues from least to greatest, so one must sort them before selecting the appropriate noise eigenvectors. The sorting method was shown in the previous example. The noise subspace is then given by $\text{EN} = \text{E}(:, \text{index}(1:M - D))$. The MATLAB code for this example demonstrates the sorting process).

Solution We can generate the 100 time samples of the noise and signal as indicated earlier. After finding the array correlation matrix $\hat{R}_{xx}$, we can use the `eig()` command in MATLAB to find the eigenvectors and corresponding eigenvalues. The eigenvalues are given by $\lambda_1 = .08$, $\lambda_2 = .09$, $\lambda_3 = .12$, $\lambda_4 = .13$, $\lambda_5 = 2.97$, and $\lambda_6 = 9$.

Applying this information to Eq. (7.35), we can plot the angular spectra in Fig. 7.9.

It is clear from the last example that the resolution of the MUSIC algorithm begins to diminish as we have to estimate the correlation matrices by time averages so that we have $\hat{R}_{xx} = \bar{A} \hat{R}_{ss} \bar{A}^H + \bar{A} \hat{R}_{sn} + \hat{R}_{ns} \bar{A}^H + \hat{R}_{nn}$.

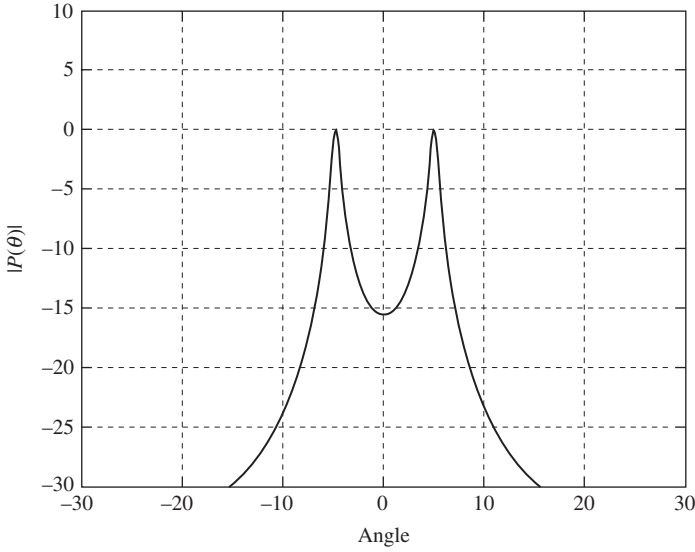


Figure 7.9 MUSIC pseudospectrum using time averages for $\theta_1 = -5^\circ$, $\theta_2 = 5^\circ$.

7.3.8 Root-MUSIC AOA estimate

The MUSIC algorithm in general can apply to any arbitrary array regardless of the position of the array elements. *Root-MUSIC* implies that the MUSIC algorithm is reduced to finding roots of a polynomial as opposed to merely plotting the pseudospectrum or searching for peaks in the pseudospectrum. Barabell [12] simplified the MUSIC algorithm for the case where the antenna is a ULA. Recalling that the MUSIC pseudospectrum is given by

$$P_{MU}(\theta) = \frac{1}{|\tilde{a}(\theta)^H \bar{E}_N \bar{E}_N^H \tilde{a}(\theta)|} \quad (7.37)$$

One can simplify the denominator expression by defining the matrix $\bar{C} = \bar{E}_N \bar{E}_N^H$ which is Hermitian. This leads to the root-MUSIC expression

$$P_{RMU}(\theta) = \frac{1}{|\tilde{a}(\theta)^H \bar{C} \tilde{a}(\theta)|} \quad (7.38)$$

If we have an ULA, the m th element of the array steering vector is given by

$$a_m(\theta) = e^{jkd(m-1)\sin\theta} \quad m = 1, 2, \dots, M \quad (7.39)$$

The denominator argument in Eq. (7.38) can be written as

$$\begin{aligned}\bar{a}(\theta)^H \bar{C} \bar{a}(\theta) &= \sum_{m=1}^M \sum_{n=1}^M e^{-jkd(m-1)\sin\theta} C_{mn} e^{jkd(n-1)\sin\theta} \\ &= \sum_{\ell=-M+1}^{M-1} c_{\ell} e^{jkd\ell\sin\theta}\end{aligned}\quad (7.40)$$

where c_{ℓ} is the sum of the diagonal elements of $\bar{C}$ along the ℓ th diagonal such that

$$c_{\ell} = \sum_{n-m=\ell} C_{mn} \quad (7.41)$$

It should be noted that the matrix $\bar{C}$ has off-diagonal sums such that $c_0 > |c_{\ell}|$ for $\ell \neq 0$. Thus the sum of off-diagonal elements is always less than the sum of the main diagonal elements. In addition, $c_{\ell} = c_{-\ell}^*$. For a 6×6 matrix we have 11 diagonals ranging from diagonal numbers $\ell = -5, -4, \dots, 0, \dots, 4, 5$. The lower left diagonal is represented by $\ell = -5$ whereas the upper right diagonal is represented by $\ell = 5$. The c_{ℓ} coefficients are calculated by $c_{-5} = C_{61}$, $c_{-4} = C_{51} + C_{62}$, $c_{-3} = C_{41} + C_{52} + C_{63}$, and so on.

We can simplify Eq. (7.40) to be in the form of a polynomial whose coefficients are c_{ℓ} . Thus

$$D(z) = \sum_{\ell=-M+1}^{M-1} c_{\ell} z^{\ell} \quad (7.42)$$

where $z = e^{-jkd\sin\theta}$

The roots of $D(z)$ that lie closest to the unit circle correspond to the poles of the MUSIC pseudospectrum. Thus, this technique is called *root-MUSIC*. The polynomial of Eq. (7.42) is of order $2(M-1)$ and thus has roots of $z_1, z_2, \dots, z_{2(M-1)}$. Each root can be complex and using polar notation can be written as

$$z_i = |z_i| e^{j\arg(z_i)} \quad i = 1, 2, \dots, 2(M-1) \quad (7.43)$$

where $\arg(z_i)$ is the phase angle of z_i .

Exact zeros in $D(z)$ exist when the root magnitudes $|z_i| = 1$. One can calculate the AOA by comparing $e^{j\arg(z_i)}$ to $e^{jkd\sin\theta_i}$ to get

$$\theta_i = -\sin^{-1} \left(\frac{1}{kd} \arg(z_i) \right) \quad (7.44)$$

Example 7.13 Repeat Example 7.12 by changing the noise variance to be $\sigma_n^2 = .3$. Change the angles of arrival to $\theta_1 = -4^\circ$ and $\theta_2 = 8^\circ$. Let the array be reduced to having only $M = 4$ elements. Approximate the correlation matrices by time averaging over $K = 300$ data points. Superimpose the plots of the pseudospectrum and the roots from root-MUSIC and compare.

Solution One can modify the MATLAB program with the new variables so that the 4-element array produces a 4×4 matrix $\bar{C}$ as defined previously. Thus $\bar{C}$ is given by

$$\bar{C} = \begin{bmatrix} .305 & -.388 + .033i & -.092 + .028i & .216 - .064i \\ -.388 - .033i & .6862 & -.225 + .019i & -.11 + .011i \\ -.092 - .028i & -.225 - .019i & .6732 & -.396 + .046i \\ .216 + .064i & -.11 - .011i & -.396 - .046i & .335 \end{bmatrix}$$

The root-MUSIC polynomial coefficients are given by the sums along the $2M - 1$ diagonals. Thus

$$c = .216 + .065i, -.203 - .039i, -1.01 - .099i, 2.0, \\ -1.01 + .099i, -.203 + .039i, .216 - .065i$$

We can use the root command in MATLAB to find the roots and then solve for the magnitude and angles of the $2(M - 1) = 6$ roots. We can plot the location of all 6 roots showing which roots are closest to the unit circle as shown in Fig. 7.10. It is clear that only the four on the right side of the y axis are nearest to the unit circle and are close to the expected angles of arrival.

We can choose the four roots closest to the unit circle and replot them along with the MUSIC pseudospectrum in Fig. 7.11.

The roots found with root-MUSIC earlier do not exactly reflect the actual location of the angles of arrival of $\theta_1 = -4^\circ$ and $\theta_2 = 8^\circ$ but they do indicate two angles of arrival. The roots themselves show the existence of an angle of arrival at near 8° which is not obvious from the plot of the MUSIC pseudospectrum. The error in locating the correct root locations owes to the fact that the incoming signals are partially correlated, that we approximated the correlation matrix by time averaging, and that the S/N ratio is relatively low. One must exert care in exercising the use of root-MUSIC by knowing the assumptions and conditions under which the calculations are made.

It is interesting to note that the polynomial $D(z)$ is a self-reciprocal polynomial such that $D(z) = D^*(z)$. The roots of the polynomial $D(z)$ are in reciprocal pairs meaning that $z_1 = \frac{1}{z_2^*}, z_3 = \frac{1}{z_4^*}, \dots, z_{2M-3} = \frac{1}{z_{2M-2}^*}$.

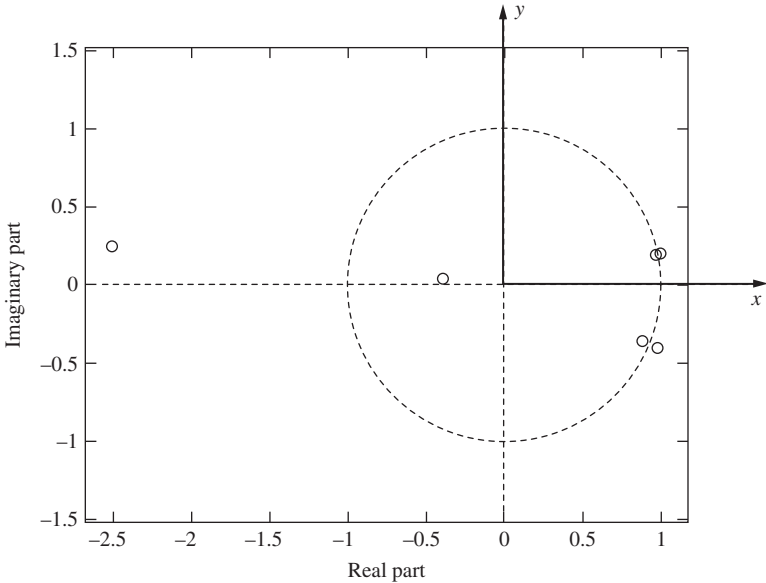


Figure 7.10 All 6 roots in cartesian coordinates.

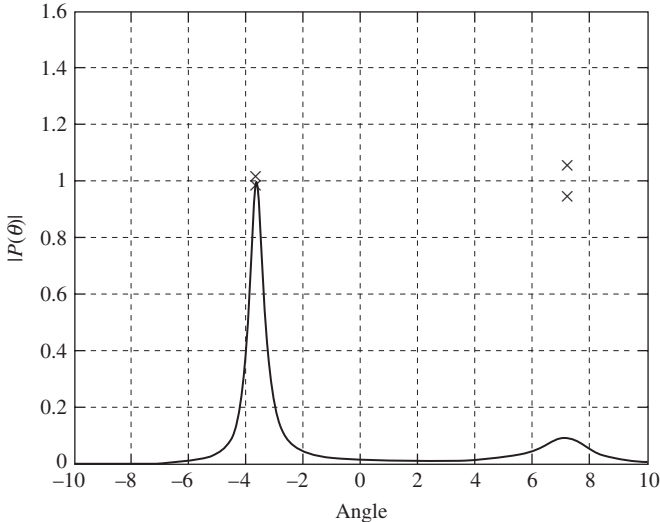


Figure 7.11 MUSIC pseudospectrum and roots found with root-MUSIC for $\theta_1 = -4^\circ$ and $\theta_2 = 8^\circ$.

Due to the self-reciprocal symmetry of $D(z)$ ([4]), one can factor $D(z)$ using the Fejér method such that

$$D(z) = p(z)p^*(1/z^*) \quad (7.45)$$

Under these conditions it is sufficient to solve for the roots of the $p(z)$ polynomial of degree $M - 1$. The roots of $p(z)$ are on or within the unit circle while the roots of $p^*(1/z^*)$ are on or without the unit circle.

One method has been proposed by Ren and Willis [19] to reduce the order of the polynomial $D(z)$ and thus reduce the computational burden of finding the roots.

The polynomial rooting method can also be applied to the Capon algorithm where we substitute $\bar{C} = \hat{R}_{rr}^{-1}$ for $\bar{C} = \bar{E}_N \bar{E}_N^H$. However, since the accuracy of the Capon estimation algorithm is much less than the MUSIC approach, the root finding also suffers a loss in accuracy.

The same principles applied to root-MUSIC can also be applied to the Min-Norm method to create a root-Min-Norm solution. We can repeat Eq. (7.32)

$$P_{RMN}(\theta) = \frac{1}{|\bar{a}(\theta)^H \bar{C} \bar{u}_1|^2} \quad (7.46)$$

where $\bar{u}_1$ = Cartesian basis vector (first column of the $M \times M$ identity matrix)

$$= [1 \ 0 \ 0 \ \dots \ 0]^T$$

$$\bar{C} = \bar{E}_N \bar{E}_N^H = \text{a } M \times M \text{ Hermitian matrix}$$

$$\bar{E}_N = \text{subspace of } M - D \text{ noise eigenvectors}$$

$$\bar{a}(\theta) = \text{array steering vector}$$

The product of the Cartesian basis vector and the Hermitian matrix results in creating a column vector composed of the first row of the matrix $\bar{C}$. The column vector based upon the first column of $\bar{C}$ becomes $\bar{c}_1 = [C_{11} \ C_{12} \ \dots \ C_{1M}]^T$ where the subscript 1 indicates the first column. We can substitute this into Eq. (7.46).

$$P_{RMN}(\theta) = \frac{1}{|\bar{a}(\theta)^H \bar{c}_1|^2} = \frac{1}{\bar{a}(\theta)^H \bar{c}_1 \bar{c}_1^H \bar{a}(\theta)} \quad (7.47)$$

In a similar fashion to Eq. (7.42), we can create a polynomial from the denominator of Eq. (7.47) given by

$$D(z) = \sum_{\ell=-M+1}^{M-1} c_\ell z^\ell \quad (7.48)$$

The coefficients c_ℓ are again the sums of the $2M - 1$ matrix diagonals of $\bar{c}_1 \bar{c}_1^H$.

Example 7.14 Apply the root-Music method to the min-norm method where $\sigma_n^2 = .3, \theta_1 = -2^\circ, \theta_2 = 4^\circ$, and $M = 4$. Approximate the correlation matrices by again time averaging over $K = 300$ data points as was done in Example 7.12. Superimpose the plots of the Min-Norm pseudospectrum and the roots from root-Min-Norm and compare.

Solution The first column of the $\bar{C}$ matrix is given by

$$\bar{c}_1 = \begin{bmatrix} .19 \\ -.33 + .02i \\ -.06 + .04i \\ .2 - .05i \end{bmatrix}$$

We can calculate the matrix $\bar{c}_1 \bar{c}_1^H$ and find the polynomial coefficients by summing along the diagonals.

$$c = .04 - .01i, -.08 + .02i, -.06 - .01i, .1937, -.06 + .01i, -.08 - .02i, .04 + .01i$$

We can use the root command in MATLAB to find the roots and then solve for the magnitude and angles of the $2(M - 1) = 6$ roots. We can plot the location of all 6 roots showing which roots are closest to the unit circle as shown in Fig. 7.12. We can also superimpose the closest roots onto the plot of the Min-Norm pseudospectrum as shown in Fig. 7.13.

The Min-Norm pseudospectrum has much sharper resolution than MUSIC but there is not indication of the AOA at -2° . However, the

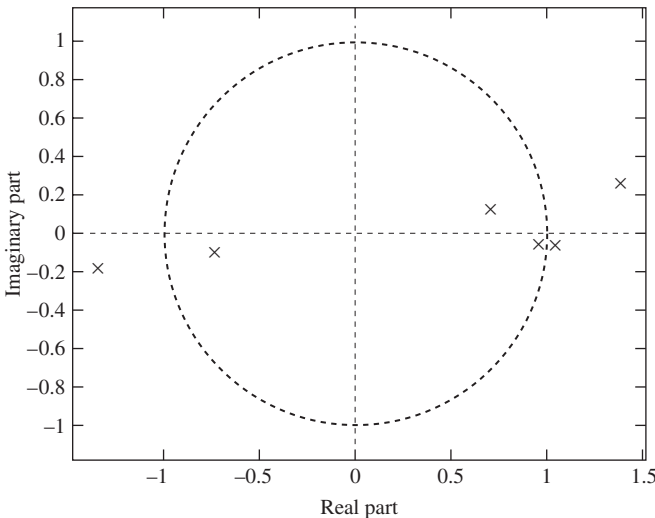


Figure 7.12 All 6 roots in cartesian coordinates.

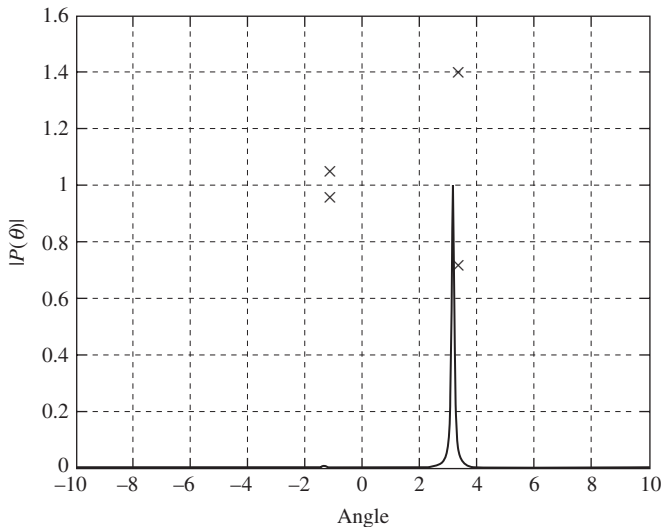


Figure 7.13 Min-norm pseudospectrum and roots found with root-Min-Norm for $\theta_1 = -2^\circ$ and $\theta_2 = 4^\circ$.

root-Min-Norm algorithm gives a fair indication of the location of both angles of arrival.

7.3.9 ESPRIT AOA estimate

ESPRIT stands for *Estimation of Signal Parameters via Rotational Invariance Techniques* and was first proposed by Roy and Kailath [20] in 1989. Useful summaries of this technique are given by both Godara [1] and Liberti and Rappaport [21]. The goal of the ESPRIT technique is to exploit the rotational invariance in the signal subspace which is created by two arrays with a translational invariance structure. ESPRIT inherently assumes narrowband signals so that one knows the translational phase relationship between the multiple arrays to be used. As with MUSIC, ESPRIT assumes that there are $D < M$ narrow-band sources centered at the center frequency f_0 . These signal sources are assumed to be of a sufficient range so that the incident propagating field is approximately planar. The sources can be either random or deterministic and the noise is assumed to be random with zero-mean. ESPRIT assumes multiple identical arrays called *doublets*. These can be separate arrays or can be composed of subarrays of one larger array. It is important that these arrays are displaced translationally but not rotationally. An example is shown in Fig. 7.14 where a four element linear array is composed of two identical three-element subarrays or two doublets. These two subarrays are translationally

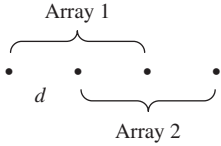


Figure 7.14 Doublet composed of two identical displaced arrays.

displaced by the distance d . Let us label these arrays as array 1 and array 2.

The signals induced on each of the arrays are given by

$$\begin{aligned}\bar{x}_1(k) &= [\bar{a}_1(\theta_1) \quad \bar{a}_1(\theta_2) \cdots \bar{a}_1(\theta_D)] \cdot \begin{bmatrix} s_1(k) \\ s_2(k) \\ \vdots \\ s_D(k) \end{bmatrix} + \bar{n}_1(k) \\ &= \bar{A}_1 \cdot \bar{s}(k) + \bar{n}_1(k)\end{aligned}\quad (7.49)$$

and

$$\begin{aligned}\bar{x}_2(k) &= \bar{A}_2 \cdot \bar{s}(k) + \bar{n}_2(k) \\ &= \bar{A}_1 \cdot \bar{\Phi} \cdot \bar{s}(k) + \bar{n}_2(k)\end{aligned}\quad (7.50)$$

where $\bar{\Phi} = \text{diag}\{e^{jkd \sin \theta_1}, e^{jkd \sin \theta_2}, \dots, e^{jkd \sin \theta_D}\}$
 = a $D \times D$ diagonal unitary matrix with phase shifts between the doublets for each AOA
 $\bar{A}_i$ = Vandermonde matrix of steering vectors for subarrays $i = 1, 2$

The complete received signal considering the contributions of both subarrays is given as

$$\bar{x}(k) = \begin{bmatrix} \bar{x}_1(k) \\ \bar{x}_2(k) \end{bmatrix} = \begin{bmatrix} \bar{A}_1 \\ \bar{A}_1 \cdot \bar{\Phi} \end{bmatrix} \cdot \bar{s}(k) + \begin{bmatrix} \bar{n}_1(k) \\ \bar{n}_2(k) \end{bmatrix}\quad (7.51)$$

We can now calculate the correlation matrix for either the complete array or for the two subarrays. The correlation matrix for the complete array is given by

$$\bar{R}_{xx} = E[\bar{x} \cdot \bar{x}^H] = \bar{A} \bar{R}_{ss} \bar{A}^H + \sigma_n^2 \bar{I}\quad (7.52)$$

whereas the correlation matrices for the two subarrays are given by

$$\bar{R}_{11} = E[\bar{x}_1 \cdot \bar{x}_1^H] = \bar{A} \bar{R}_{ss} \bar{A}^H + \sigma_n^2 \bar{I}\quad (7.53)$$

and

$$\bar{R}_{22} = E [\bar{x}_2 \cdot \bar{x}_2^H] = \bar{A} \bar{\Phi} \bar{R}_{ss} \bar{\Phi}^H \bar{A}^H + \sigma_n^2 \bar{I} \quad (7.54)$$

Each of the full rank correlation matrices given in Eq. (7.53) and (7.54) has a set of eigenvectors corresponding to the D signals present. Creating the signal subspace for the two subarrays results in the two matrices $\bar{E}_1$ and $\bar{E}_2$. Creating the signal subspace for the entire array results in one signal subspace given by $\bar{E}_x$. Because of the invariance structure of the array, $\bar{E}_x$ can be decomposed into the subspaces $\bar{E}_1$ and $\bar{E}_2$.

Both $\bar{E}_1$ and $\bar{E}_2$ are $M \times D$ matrices whose columns are composed of the D eigenvectors corresponding to the largest eigenvalues of $\bar{R}_{11}$ and $\bar{R}_{22}$. Since the arrays are translationally related, the subspaces of eigenvectors are related by a unique non-singular transformation matrix $\bar{\Psi}$ such that

$$\bar{E}_1 \bar{\Psi} = \bar{E}_2 \quad (7.55)$$

There must also exist a unique non-singular transformation matrix $\bar{T}$ such that

$$\bar{E}_1 = \bar{A} \bar{T} \quad (7.56)$$

and

$$\bar{E}_2 = \bar{A} \bar{\Phi} \bar{T} \quad (7.57)$$

By substituting Eqs. (7.55) and (7.56) into Eq. (7.57) and assuming that $\bar{A}$ is of full-rank, we can derive the relationship

$$\bar{T} \bar{\Psi} \bar{T}^{-1} = \bar{\Phi} \quad (7.58)$$

Thus, the eigenvalues of $\bar{\Psi}$ must be equal to the diagonal elements of $\bar{\Phi}$ such that $\lambda_1 = e^{jkd \sin \theta_1}$, $\lambda_2 = e^{jkd \sin \theta_2}$, $\dots$, $\lambda_D = e^{jkd \sin \theta_D}$ and the columns of $\bar{T}$ must be the eigenvectors of $\bar{\Psi}$. $\bar{\Psi}$ is a rotation operator that maps the signal subspace $\bar{E}_1$ into the signal subspace $\bar{E}_2$. One is now left with the problem of estimating the subspace rotation operator $\bar{\Psi}$ and consequently finding the eigenvalues of $\bar{\Psi}$.

If we are restricted to a finite number of measurements and we also assume that the subspaces $\bar{E}_1$ and $\bar{E}_2$ are equally noisy, we can estimate the rotation operator $\bar{\Psi}$ using the *total least-squares* (TLS) criterion. Details of the TLS criterion can be found in van Huffel and Vandewalle [22]. This procedure is outlined as follows. (see Roy and Kailath [20]).

- Estimate the array correlation matrices $\bar{R}_{11}$, $\bar{R}_{22}$ from the data samples.
- Knowing the array correlation matrices for both subarrays, one can estimate the total number of sources by the number of large eigenvalues in either $\bar{R}_{11}$ or $\bar{R}_{22}$.
- Calculate the signal subspaces $\bar{E}_1$ and $\bar{E}_2$ based upon the signal eigenvectors of $\bar{R}_{11}$ and $\bar{R}_{22}$. For ULA, one can alternatively construct the signal subspaces from the entire array signal subspace $\bar{E}_x$. $\bar{E}_x$ is an $M \times D$ matrix composed of the signal eigenvectors. $\bar{E}_1$ can be constructed by selecting the first $M/2 + 1$ rows ($(M + 1)/2 + 1$ for odd arrays) of $\bar{E}_x$. $\bar{E}_2$ can be constructed by selecting the last $M/2 + 1$ rows ($(M + 1)/2 + 1$ for odd arrays) of $\bar{E}_x$.
- Next form a $2D \times 2D$ matrix using the signal subspaces such that

$$\bar{C} = \begin{bmatrix} \bar{E}_1^H \\ \bar{E}_2^H \end{bmatrix} \quad [\bar{E}_1 \quad \bar{E}_2] = \bar{E}_C \bar{\Lambda} \bar{E}_C^H \quad (7.59)$$

where the matrix $\bar{E}_C$ is from the *eigenvalue decomposition* (EVD) of $\bar{C}$ such that $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{2D}$ and $\bar{\Lambda} = \text{diag} \{\lambda_1, \lambda_2, \dots, \lambda_{2D}\}$

- Partition $\bar{E}_C$ into four $D \times D$ submatrices such that

$$\bar{E}_C = \begin{bmatrix} \bar{E}_{11} & \bar{E}_{12} \\ \bar{E}_{21} & \bar{E}_{22} \end{bmatrix} \quad (7.60)$$

- Estimate the rotation operator $\bar{\Psi}$ by

$$\bar{\Psi} = -\bar{E}_{12} \bar{E}_{22}^{-1} \quad (7.61)$$

- Calculate the eigenvalues of $\bar{\Psi}$, $\lambda_1, \lambda_2, \dots, \lambda_D$
- Now estimate the angles of arrival, given that $\lambda_i = |\lambda_i| e^{j \arg(\lambda_i)}$

$$\theta_i = \sin^{-1} \left(\frac{\arg(\lambda_i)}{kd} \right) \quad i = 1, 2, \dots, D \quad (7.62)$$

If so desired, one can estimate the matrix of steering vectors from the signal subspace $\bar{E}_s$ and the eigenvectors of $\bar{\Psi}$ given by $\bar{E}_\psi$ such that $\hat{A} = \bar{E}_s \bar{E}_\psi$.

Example 7.15 Use the ESPRIT algorithm to predict the angles of arrival for an $M = 4$ -element array where noise variance $\sigma_n^2 = .1$. Approximate the correlation matrices by again time averaging over $K = 300$ data points as was done in Example 7.12. The angles of arrival are $\theta_1 = -5^\circ$, $\theta_2 = 10^\circ$.

Solution The signal subspace for the entire ideal ULA array correlation matrix is given by

$$\bar{E}_x = \begin{bmatrix} .78 & .41 + .1i \\ .12 - .02i & .56 \\ -.22 + .07i & .54 - .05i \\ -.51 + .25i & .46 - .1i \end{bmatrix}$$

The two subarray signal subspaces can now be found by taking the first three rows of $\bar{E}_x$ to define $\bar{E}_1$ and the last three rows of $\bar{E}_x$ to define $\bar{E}_2$

$$\bar{E}_1 = \begin{bmatrix} .78 & .41 + .1i \\ .12 - .02i & .56 \\ -.22 + .07i & .54 - .05i \end{bmatrix} \quad \bar{E}_2 = \begin{bmatrix} .12 - .02i & .56 \\ -.22 + .07i & .54 - .05i \\ -.51 + .25i & .46 - .1i \end{bmatrix}$$

Constructing the matrix of signal subspaces we get

$$\bar{C} = \begin{bmatrix} \bar{E}_1^H \\ \bar{E}_2^H \end{bmatrix} [\bar{E}_1 \quad \bar{E}_2] = \begin{bmatrix} .67 & .26 + .06i & .2 - .03i & .39 \\ .26 - .06i & .78 & -.37 + .12i & .78 - .11i \\ .2 + .03i & -.37 - .12i & .4 & -.32 - .08i \\ .39 & .78 + .11i & -.32 + .08i & .82 \end{bmatrix}$$

Performing the eigendecomposition we can construct the matrix $\bar{E}_C$ such that

$$\bar{E}_C = \begin{bmatrix} \bar{E}_{11} & \bar{E}_{12} \\ \bar{E}_{21} & \bar{E}_{22} \end{bmatrix} = \begin{bmatrix} .31 & .8 & -.44 + .1i & -.22 + .11i \\ .63 - .09i & -.16 & .45 & -.61 - .03i \\ -.26 - .05i & .57 + .1i & .75 & .15 - .11i \\ .66 & .01 + .02i & .07 - .17i & .73 \end{bmatrix}$$

We can now calculate the rotation operator $\bar{\Psi} = -\bar{E}_{12}\bar{E}_{22}^{-1}$ given the rotation matrix

$$\bar{\Psi} = -\bar{E}_{12}\bar{E}_{22}^{-1} = \begin{bmatrix} .58 - .079i & .2 - .05i \\ -.67 + .23i & .94 - .11i \end{bmatrix}$$

Next we can calculate the eigenvalues of $\bar{\Psi}$ and solve for the angles of arrival using

$$\theta_1 = \sin^{-1} \left(\frac{\arg(\lambda_1)}{kd} \right) = -4.82^\circ$$

$$\theta_2 = \sin^{-1} \left(\frac{\arg(\lambda_2)}{kd} \right) = 9.85^\circ$$

References

- Godara, L., "Application of Antenna Arrays to Mobile Communications, Part II: Beam-Forming and Direction-of-Arrival Considerations," *Proceedings of the IEEE*, Vol. 85, No. 8, pp. 1195–1245, Aug. 1997.
- Capon, J., "High-Resolution Frequency-Wavenumber Spectrum Analysis," *Proceedings of the IEEE*, Vol. 57, No. 8, pp. 1408–1418, Aug. 1969.
- Johnson, D., "The Application of Spectral Estimation Methods to Bearing Estimation Problems," *Proceedings of the IEEE*, Vol. 70, No. 9, pp. 1018–1028, Sept. 1982.
- Van Trees, H., *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*, Wiley Interscience, New York, 2002.
- Stoica, P., and R. Moses, *Introduction to Spectral Analysis*, Prentice Hall, New York, 1997.
- Shan, T.-J., M. Wax, and T. Kailath, "Spatial Smoothing for Direction-of-Arrival Estimation of Coherent Signals," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. ASSP-33, No. 4, pp. 806–811, Aug. 1985.
- Minasyan, G., "Application of High Resolution Methods to Underwater Data Processing," Ph.D. Dissertation, N.N. Andreyev Acoustics Institute, Moscow, Oct. 1994 (In Russian).
- Bartlett, M., *An Introduction to Stochastic Processes with Special References to Methods and Applications*, Cambridge University Press, New York, 1961.
- Makhoul, J., "Linear Prediction: A Tutorial Review," *Proceedings of IEEE*, Vol. 63, pp. 561–580, 1975.
- Burg, J.P., "Maximum Entropy Spectrum Analysis," Ph.D. dissertation, Dept. of Geophysics, Stanford University, Stanford CA, 1975.
- Burg, J.P., "The Relationship Between Maximum Entropy Spectra and Maximum Likelihood Spectra," *Geophysics*, Vol. 37, pp. 375–376, April 1972.
- Barabell, A., "Improving the Resolution of Eigenstructure-Based Direction-Finding Algorithms," *Proceedings of ICASSP*, Boston, MA, pp. 336–339, 1983.
- Pisarenko, V.F., "The Retrieval of Harmonics from a Covariance Function," *Geophysical Journal of the Royal Astronomical Society*, Vol. 33 pp. 347–366, 1973.
- Johnson D., and D. Dudgeon, *Array Signal Processing Concepts and Techniques*, Prentice Hall Signal Processing Series, New York, 1993.
- Reddi, S.S., "Multiple Source Location—A Digital Approach," *IEEE Transactions on AES*, Vol. 15, No. 1, Jan. 1979.
- Kumaresan, R., and D. Tufts, "Estimating the Angles of Arrival of Multiple Plane Waves," *IEEE Transactions on AES*, Vol. AES-19, pp. 134–139, 1983.
- Ermolaev, V., and A. Gershman, "Fast Algorithm for Minimum-Norm Direction-of-Arrival Estimation," *IEEE Transactions on Signal Processing*, Vol. 42, No. 9, Sept. 1994.
- Schmidt, R., "Multiple Emitter Location and Signal Parameter Estimation," *IEEE Transactions on Antenna. Propagation.*, Vol. AP-34, No. 2, pp. 276–280, March 1986.
- Ren, Q., and A. Willis, "Fast root-MUSIC Algorithm," *IEE Electronics Letters*, Vol. 33, No. 6, pp. 450–451, March 1997.
- Roy, R., and T. Kailath, "ESPRIT—Estimation of Signal Parameters Via Rotational Invariance Techniques," *IEEE Transactions on ASSP*, Vol. 37, No. 7, pp. 984–995, July 1989.
- Liberti, J., and T. Rappaport, *Smart Antennas for Wireless Communications*, Prentice Hall, New York, 1999.
- van Huffel, S., and J. Vandewalle, "The Total Least Squares Problem: Computational Aspects and Analysis," SIAM, Philadelphia, PA, 1991.

Problems

- 7.1 For the two matrices $\bar{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ $\bar{B} = \begin{bmatrix} 1 & 1 \\ 5 & 2 \end{bmatrix}$
- What is $\text{Trace}(\bar{A})$ and $\text{Trace}(\bar{B})$
 - Show that $(\bar{A} \cdot \bar{B})^T = \bar{B}^T \cdot \bar{A}^T$

7.2 For the two matrices $\bar{A} = \begin{bmatrix} 1 & 2j \\ 2j & 2 \end{bmatrix}$ $\bar{B} = \begin{bmatrix} j & 1 \\ 1 & 2j \end{bmatrix}$ show that $(\bar{A} \cdot \bar{B})^H = \bar{B}^H \cdot \bar{A}^H$

7.3 For $\bar{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$
 (a) Find the $\bar{A}^{-1}$ by hand
 (b) Find the $\bar{A}^{-1}$ using MATLAB

7.4 For $\bar{A} = \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix}$
 (a) Solve for all eigenvalues by hand using Eq. (7.16)
 (b) Solve for all eigenvectors by hand using Eq. (7.17)
 (c) Solve for all eigenvectors and eigenvalues using MATLAB

7.5 Repeat Prob. 4 for $\bar{A} = \begin{bmatrix} 2 & 4 \\ .25 & 2 \end{bmatrix}$

7.6 For $\bar{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$
 (a) What are all eigenvalues and eigenvectors using MATLAB?
 (b) Which eigenvalues are associated with which eigenvectors?
 (c) Use Eq. (7.17), and by hand, prove that the first eigenvalue and the first eigenvector satisfy this equation.

7.7 An $N = 3$ -element array exists with element spacing $d = \lambda/2$. One arriving signal is defined as $s_1(k) = .1, .2, .3$; ($k = 1, 2, 3$) arrives at the angle $\theta_1 = 0^\circ$. The other signal is defined as $s_2(k) = .3, .4, .5$; ($k = 1, 2, 3$) arrives at the angle $\theta_2 = 30^\circ$. The noise has a standard deviation $\sigma = .1$. Use Eqs. (7.15), (7.16), and (7.17). Use MATLAB.

- What are the array steering vectors $\bar{a}(\theta_1)$, $\bar{a}(\theta_2)$?
- What is the matrix of steering vectors $\bar{A}$?
- What are the correlation matrices $\bar{R}_{ss}$, $\bar{R}_{nn}$, $\bar{R}_{xx}$?
- What are the eigenvalues and eigenvectors of $\bar{R}_{xx}$?

7.8 Repeat parts *c* and *d* of Prob. 7 using estimates on the correlation matrices ($\hat{R}_{ss}$, $\hat{R}_{nn}$, $\hat{R}_{xx}$). Allow the noise to be defined in MATLAB as $n = 0.1 \cdot \text{randn}(3, 3)$. Thus the noise is Gaussian distributed but we only can work with three time samples. Are the eigenvalues and eigenvectors similar to those found in Prob. 7. Why or why not?

7.9 Plot the normalized Bartlett pseudospectrum $P_B(\theta)$ using MATLAB, for the case where $M = 7$, $d = \lambda/2$, $\sigma_n^2 = .2$, and $\theta_1 = 3^\circ$, $\theta_2 = -3^\circ$. Set the vertical scale to be -30 to 5 dB and horizontal scale to be -15° to 15° . Allow the signals s_1 and s_2 to be uncorrelated resulting in $\bar{R}_{ss} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

7.10 Plot the normalized Capon pseudospectrum $P_C(\theta)$ using MATLAB, for the case where $M = 7$, $d = \lambda/2$, $\sigma_n^2 = .2$, and $\theta_1 = 3^\circ$, $\theta_2 = -3^\circ$. Set the vertical scale to be -30 to 5 dB and horizontal scale to be -15° to 15° . Allow the signals s_1 and s_2 to be uncorrelated resulting in $\bar{R}_{ss} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

7.11 Plot the normalized linear prediction pseudospectrum $P_{LP4}(\theta)$ using MATLAB, for the case where $M = 7$, $d = \lambda/2$, $\sigma_n^2 = .2$, and $\theta_1 = 3^\circ$, $\theta_2 = -3^\circ$. Set the vertical scale to be -30 to 5 dB and horizontal scale to be -15° to 15° . Allow the signals s_1 and s_2 to be uncorrelated resulting in $\bar{R}_{ss} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

7.12 Plot the normalized maximum entropy pseudospectrum $P_{ME4}(\theta)$ using MATLAB, for the case where $M = 7$, $d = \lambda/2$, $\sigma_n^2 = .2$, and $\theta_1 = 3^\circ$, $\theta_2 = -3^\circ$. Set the vertical scale to be -30 to 5 dB and horizontal scale to be -15° to 15° . Allow the signals s_1 and s_2 to be uncorrelated resulting in $\bar{R}_{ss} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

7.13 For the Pisarenko harmonic decomposition pseudospectrum $P_{PHD}(\theta)$ where $M = 7$, $d = \lambda/2$, $\sigma_n^2 = .2$, and $\theta_1 = 3^\circ$, $\theta_2 = -3^\circ$. Set the vertical scale to be -30 to 5 dB and horizontal scale to be -15° to 15° . Allow the signals s_1 and s_2 to be uncorrelated resulting in $\bar{R}_{ss} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

- What is the smallest eigenvalue?
- What is the eigenvector associated with that eigenvalue?
- Plot the normalized pseudospectrum using MATLAB

7.14 For the min-norm decomposition pseudospectrum $P_{MN}(\theta)$ where $M = 7$, $d = \lambda/2$, $\sigma_n^2 = .2$, and $\theta_1 = 3^\circ$, $\theta_2 = -3^\circ$, set the vertical scale to be -30 to 5 dB and horizontal scale to be -15° to 15° . Allow the signals s_1 and s_2 to be uncorrelated resulting in $\bar{R}_{ss} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

- What is the signal subspace $\bar{E}_S$?
- What is the noise subspace $\bar{E}_N$?
- Plot the normalized pseudospectrum using MATLAB

7.15 For the MUSIC decomposition pseudospectrum $P_{MUSIC}(\theta)$ where $M = 7$, $d = \lambda/2$, $\sigma_n^2 = .2$, and $\theta_1 = 3^\circ$, $\theta_2 = -3^\circ$, set the vertical scale to be -30 to 5 dB and horizontal scale to be -15° to 15° . Allow the signals s_1 and s_2 to be uncorrelated resulting in $\bar{R}_{ss} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

- What is the signal subspace $\bar{E}_S$?
- What is the noise subspace $\bar{E}_N$?
- Plot the normalized pseudospectrum using MATLAB

7.16 Repeat Example 7.12 using $M = 7$, $d = \lambda/2$, $\sigma_n^2 = .2$, and $\theta_1 = 3^\circ$, $\theta_2 = -3^\circ$. Set the vertical scale to be -30 to 5 dB and horizontal scale to be -15° to 15° .

7.17 For an $M = 3$ -element array with $d = \lambda/2$, $\sigma_n^2 = .2$, and $\theta_1 = 3^\circ$, $\theta_2 = -3^\circ$ and given that the signals s_1 and s_2 are uncorrelated, apply the root-MUSIC method

- What is the matrix $\bar{C}$?
- What are the coefficients c_ℓ using Eq. (7.47)?
- What are the roots z_i using Eq. (7.49)?
- What are the angles θ_i using Eq. (7.50)?

7.18 For an $M = 3$ -element array with $d = \lambda/2$, $\sigma_n^2 = .2$, and $\theta_1 = 3^\circ$, $\theta_2 = -3^\circ$ and given that the signals s_1 and s_2 are uncorrelated, apply the root-Min-Norm method

- (a) What is the matrix $\bar{C}$?
- (b) What is the first column $\bar{c}_1$?
- (c) What are the coefficients c_ℓ using Eq. (7.47)?
- (d) What are the roots z_i using Eq. (7.49)?
- (e) What are the angles θ_i using Eq. (7.50)?

7.19 For an $M = 4$ -element array with $d = \lambda/2$, $\sigma_n^2 = .2$, and $\theta_1 = 3^\circ$, $\theta_2 = -3^\circ$ and given that the signals s_1 and s_2 are uncorrelated, approximate the correlation matrices by time averaging over $K = 300$ data points as was done in Example 7.12. Apply the ESPRIT method

- (a) What are the correlation matrices $\bar{R}_{11}$, $\bar{R}_{22}$?
- (b) Find the signal subspaces $\bar{E}_1$, $\bar{E}_2$?
- (c) What is the matrix $\bar{E}_C$?
- (d) What is the rotation operator $\bar{\Psi} = -\bar{E}_{12}\bar{E}_{22}^{-1}$?
- (e) What are the eigenvalues of $\bar{\Psi}$?
- (f) What are the predicted angles of arrival?

This page intentionally left blank

Smart Antennas

8.1 Introduction

Traditional array antennas, where the main beam is steered to directions of interest, are called *phased arrays*, *beamsteered arrays*, or *scanned arrays*. The beam is steered via phase shifters and in the past these phase shifters were often implemented at RF frequencies. This general approach to phase shifting has been referred to as electronic beamsteering because of the attempt to change the phase of the current directly at each antenna element.

Modern beamsteered array antennas, where the pattern is shaped according to certain optimum criteria, are called *smart antennas*. Smart antennas have alternatively been called *digital beamformed* (DBF) arrays or *adaptive arrays* (when adaptive algorithms are employed). The term *smart* implies the use of signal processing in order to shape the beam pattern according to certain conditions. For an array to be *smart* implies sophistication beyond merely steering the beam to a direction of interest. Smart essentially means computer control of the antenna performance. Smart antennas hold the promise for improved radar systems, improved system capacities with mobile wireless, and improved wireless communications through the implementation of *space division multiple access* (SDMA).

Smart antenna patterns are controlled via algorithms based upon certain criteria. These criteria could be maximizing the *signal-to-interference ratio* (SIR), minimizing the variance, minimizing the *mean-square error* (MSE), steering toward a signal of interest, nulling the interfering signals, or tracking a moving emitter to name a few. The implementation of these algorithms can be performed *electronically* through analog devices but it is generally more easily performed using digital signal processing. This requires that the array outputs be

digitized through the use of an A/D converter. This digitization can be performed at either IF or baseband frequencies. Since an antenna pattern (or beam) is formed by digital signal processing, this process is often referred to as *digital beamforming*. Figure 8.1 contrasts a traditional electronically steered array with a DBF array or smart antenna.

When the algorithms used are adaptive algorithms, this process is referred to as *adaptive beamforming*. Adaptive beamforming is a subcategory under the more general subject of digital beamforming. Digital beamforming has been applied to radar systems [1–5], sonar systems [6], and communications systems [7] to name a few. The chief advantage of digital beamforming is that phase shifting and array weighting can be performed on the digitized data rather than by being implemented in hardware. On receive, the beam is formed in the data processing rather than literally being forming in space. The digital beamforming method cannot be strictly called electronic steering since no effort is made to directly shift the phase of the antenna element currents. Rather, the *phase shifting* is computationally performed on the digitized signal. If the parameters of operation are changed or the detection criteria are modified, the beamforming can be changed by simply changing an algorithm rather than by replacing hardware.

Adaptive beamforming is generally the more useful and effective beamforming solution because the digital beamformer merely consists of an algorithm which dynamically *optimizes* the array pattern according to the changing electromagnetic environment.

Conventional array *static* processing systems are subject to degradation by various causes. The array SNR can be severely degraded by the presence of unwanted interfering signals, electronic countermeasures,

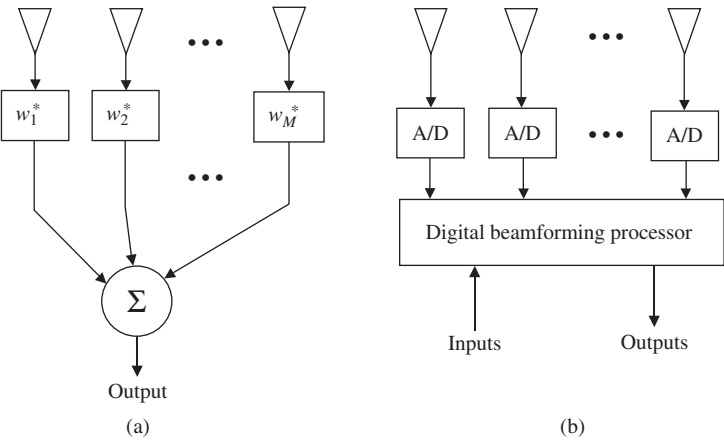


Figure 8.1 (a) Analog beamforming, (b) Digital beamforming.

clutter returns, reverberation returns (in acoustics), or multipath interference and fading. An *adaptive* array system consists of the antenna array elements terminated in an adaptive processor which is designed to specifically maximize certain criteria. As the emitters move or change, the adaptive array updates and compensates iteratively in order to track the changing environment. Many current modern radar systems still rely on older electronic scanning technologies. Recent efforts are being exerted to modify radar systems to include digital beamforming and adaptive beamforming techniques [4]. While current modern mobile base stations tend to use older fixed beam technologies to satisfy SDMA, they also would benefit from the use of modern adaptive methods and thereby increase system capacities [8].

This chapter will review the historical development of digital beamforming and adaptive arrays, explain some basic DBF principles, and will cover the more popular DBF methods which include adaptive methods.

8.2 The Historical Development of *Smart Antennas*

“The development of adaptive arrays began in the late 50s and has been more than four decades in the making.” The word *adaptive array* was first coined by Van Atta [9], in 1959, to describe a *self-phased* array. Self-phased arrays merely reflected all incident signals back in the direction of arrival by using clever phasing schemes based upon phase conjugation. Because of the redirection of incident planewaves, these arrays were called *retrodirective arrays* (see an in-depth explanation of retrodirective arrays in Chap. 4, Sect. 4.9). Self-phased arrays are instantaneously adaptive arrays since they essentially reflect the incident signal in a similar fashion to the classic corner reflector.

Phase-locked loop (PLL) systems were incorporated into arrays in the 60s in an effort to construct better retrodirective arrays since it was assumed that retrodirection was the best approach [10]. PLLs still are used in single beam scanning systems [11].

Adaptive *sidelobe cancellation* (SLC) was first proposed by Howells [12, 13] in 1959. This technique allowed for interference nulling, thus raising the SIR. The Howells SLC was the first adaptive scheme which allowed for automatic interference nulling. By maximizing the generalized *signal-to-noise ratio* (SNR), Applebaum developed the algorithm governing adaptive interference cancellation [14, 15]. His algorithm became known as the *Howells-Applebaum algorithm*. At the same time, through the use of *least mean squares* (LMS), Widrow and others applied self-training to adaptive arrays [16, 17]. The Howells-Applebaum and *Widrow algorithms* both are steepest-descent/gradient-search

methods. The Howells-Applebaum and Widrow algorithms both converge to the optimum Wiener solution. The convergence of these methods is dependent upon the eigenvalue spread [18] such that larger spreads require longer convergence times. The convergence time constant is given by [19]

$$\tau_i = \frac{1}{2\mu\lambda_i} \quad (8.1)$$

Where μ = gradient step size
 λ_i = i th eigenvalue.

The eigenvalues are derived from the correlation matrix where the largest eigenvalues correspond to the strongest signals and the smallest eigenvalues correspond to the weakest signals or noise. The larger eigenvalue spreads result in longer convergence times. In the case of SLC, the weakest interfering signals are cancelled last as can be seen from Eq. (8.1).

Since the convergence of SLC algorithm was slow for large eigenvalue spreads, Reed, Mallett, and Brennen developed the direct *sample matrix inversion* (SMI) technique in 1974 [20].

The next great complementary advance in adaptive arrays came with the application of spectral estimation methods to array processing. (Many of these spectral estimation methods are discussed at length in Chap. 7.) The spectral estimation methods, achieving higher angular resolutions, have come to be known as *superresolution* algorithms [18]. Capon, in 1969, used a *maximum likelihood* (ML) method to solve for the *minimum variance distortionless response* (MVDR) of an array. His solution maximizes the SIR [21]. Additionally, the linear predictive method was used to minimize the mean-squared prediction error leading to optimum array weights. The array weights are dependent on the array correlation matrix and are given in Eq. (7.22) [22]. In 1972, Burg applied the maximum entropy method to spectral estimation and his technique was soon adapted to array signal processing [23, 24]. In 1973, Pisarenko developed the harmonic decomposition technique based upon minimizing the MSE under the constraint that the norm of the weight vector be equal to unity [25]. The minimum-norm (min-norm) method was developed by Reddi [26] in 1979 and Kumaresan and Tufts [27] in 1983. The min-norm algorithm optimizes the weight vector by solving the optimization problem where the weight vector is orthogonal to the signal eigenvector subspace. This too is a spectral estimation problem applied to array signal processing. The now famous MUSIC algorithm was developed by Schmidt in 1986 [28]. MUSIC is a spectral estimation based algorithm which exploits the orthogonality of the noise subspace with the array correlation matrix. The estimation of signal parameters

via rotational invariance (ESPRIT) technique was first proposed by Roy and Kailath [29] in 1989. The goal of ESPRIT is to exploit the rotational invariance in the signal subspace which is created by two arrays with a translational invariance structure.

This latter group of adaptive methods can be considered to be part of the array *superresolution* methods allowing the user to achieve higher resolutions than that allowed by the beamwidth of the array. The price of increased resolution comes at the cost of greater computational intensity.

Let us review some of the basics for fixed and adaptive algorithms.

8.3 Fixed Weight Beamforming Basics

8.3.1 Maximum signal-to-interference ratio

One criterion which can be applied to enhancing the received signal and minimizing the interfering signals is based upon maximizing the SIR [7, 30]. Before we engage in the rigorous derivation of the SIR optimization, let us review a heuristic approach which was used in Chap. 4. It is intuitive that if we can cancel all interference by placing nulls at their angles of arrival, we will automatically maximize the SIR.

Repeating the basic development of Chap. 4, let us assume an $N = 3$ -element array with one fixed known desired source and two fixed undesired interferers. All signals are assumed to operate at the same carrier frequency. Let us assume a three-element array with the desired signal and interferers as shown in Fig. 8.2.

The array vector is given by

$$\vec{a} = [e^{-jkd \sin \theta} \quad 1 \quad e^{jkd \sin \theta}]^T \tag{8.2}$$

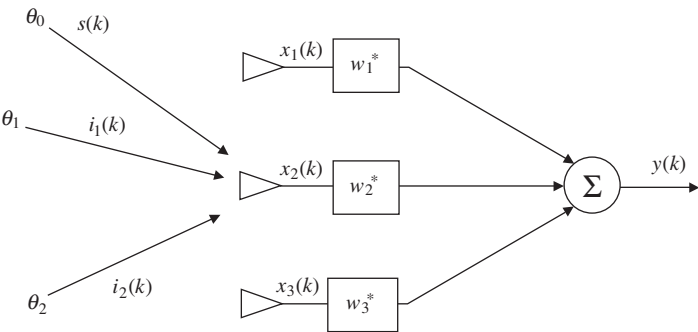


Figure 8.2 Three-element array with desired and interfering signals

The, as yet to be determined, array weights for optimization are given by

$$\bar{w}^H = [w_1 \quad w_2 \quad w_3] \quad (8.3)$$

Therefore, the general total array output is given as

$$y = \bar{w}^H \cdot \bar{a} = w_1 e^{-jkd \sin \theta} + w_2 + w_3 e^{jkd \sin \theta} \quad (8.4)$$

The array output for the desired signal will be designated by y_s whereas the array output for the interfering or undesired signals will be designated by y_1 and y_2 . Since there are three unknown weights, there must be three conditions satisfied.

$$\text{Condition 1: } y_s = \bar{w}^H \cdot \bar{a}_0 = w_1 e^{-jkd \sin \theta_0} + w_2 + w_3 e^{jkd \sin \theta_0} = 1$$

$$\text{Condition 2: } y_1 = \bar{w}^H \cdot \bar{a}_1 = w_1 e^{-jkd \sin \theta_1} + w_2 + w_3 e^{jkd \sin \theta_1} = 0$$

$$\text{Condition 3: } y_2 = \bar{w}^H \cdot \bar{a}_2 = w_1 e^{-jkd \sin \theta_2} + w_2 + w_3 e^{jkd \sin \theta_2} = 0$$

Condition 1 demands that $y_s = 1$ for the desired signal, thus allowing the desired signal to be received without modification. Conditions 2 and 3 reject the undesired interfering signals. These conditions can be recast in matrix form as

$$\bar{w}^H \cdot \bar{A} = \bar{u}_1^T \quad (8.5)$$

where $\bar{A} = [\bar{a}_0 \quad \bar{a}_1 \quad \bar{a}_2] =$ matrix of steering vectors

$\bar{u}_1 = [1 \quad 0 \quad \dots \quad 0]^T =$ Cartesian basis vector

One can invert the matrix to find the required complex weights w_1, w_2 , and w_3 by using

$$\bar{w}^H = \bar{u}_1^T \cdot \bar{A}^{-1} \quad (8.6)$$

As an example, if the desired signal is arriving from $\theta_0 = 0^\circ$ while $\theta_1 = -45^\circ$ and $\theta_2 = 60^\circ$, the necessary weights can be calculated to be

$$\begin{bmatrix} w_1^* \\ w_2^* \\ w_3^* \end{bmatrix} = \begin{bmatrix} .28 - .07i \\ .45 \\ .28 + .07i \end{bmatrix} \quad (8.7)$$

The array factor is shown plotted in Fig. 8.3.

The Cartesian basis vector in Eq. (8.6) indicates that the array weights are taken from the first row of $\bar{A}^{-1}$.

The previous development is predicated on the fact that the desired signal and the total of the interfering signals make $\bar{A}$ an invertible square matrix. $\bar{A}$ must be an $N \times N$ matrix with N -array elements and N -arriving signals. In the case where the number of interferers

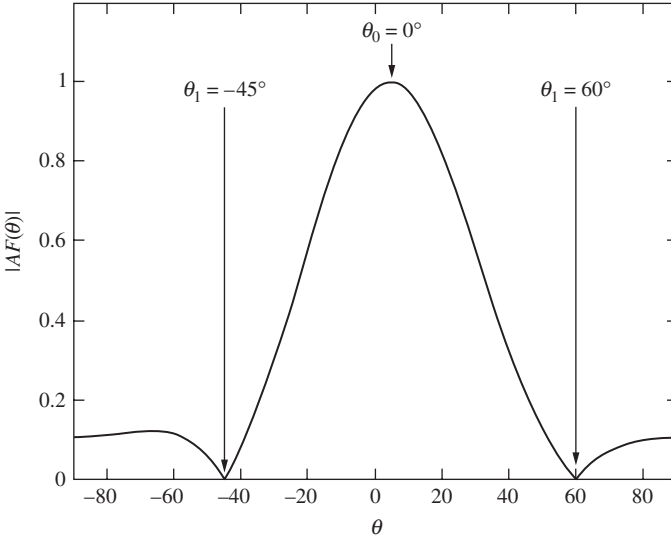


Figure 8.3 Sidelobe cancellation.

is less than $M - 1$, Godara [19] has provided an equation which gives an estimate of the weights. However, his formulation requires noise be added in the system because the matrix inversion will be singular otherwise. Using the Godara method we have

$$\bar{w}^H = \bar{u}_1^T \cdot \bar{A}^H (\bar{A} \cdot \bar{A}^H + \sigma_n^2 \bar{I})^{-1} \quad (8.8)$$

where $\bar{u}_1^T$ is the Cartesian basis vector whose length equals the total number of sources.

Example 8.1 For an $M = 5$ -element array with $d = \lambda/2$, the desired signal arriving at $\theta = 0^\circ$ and one interferer arrives at -15° while the other interferer arrives at $+25^\circ$. If the noise variance is $\sigma_n^2 = .001$, use the array weight estimation found in Eq. (8.8) to find the weights and plot.

Solution The problem is solved in MATLAB using `sa_ex8_1.m`. The matrix of steering vectors is given as

$$\bar{A} = [\bar{a}_0 \quad \bar{a}_1 \quad \bar{a}_2]$$

where

$$\begin{aligned} \bar{a}_0 &= [1 \quad 1 \quad 1 \quad 1 \quad 1]^T \\ \bar{a}_n &= [e^{-j2\pi \sin \theta_n} \quad e^{-j\pi \sin \theta_n} \quad 1 \quad e^{j\pi \sin \theta_n} \quad e^{j2\pi \sin \theta_n}]^T \quad n = 1, 2 \end{aligned}$$

Since only three sources are present, $\bar{u}_1 = [1 \quad 0 \quad 0]^T$

Substituting into Eq. (8.8)

$$\bar{w}^H = \bar{u}_1^T \cdot \bar{A}^H (\bar{A} \cdot \bar{A}^H + \sigma_n^2 \bar{I})^{-1} = \begin{bmatrix} .26 + .11i \\ .17 + .08i \\ .13 \\ .17 - .08i \\ .26 - .11i \end{bmatrix}^T$$

The plot of the array factor is shown in Fig. 8.4. The advantage of the Godara method is that the total number of sources can be less than the number of array elements.

This basic sidelobe canceling scheme works through an intuitive application of the array steering vector for the desired signal and interfering signals. However, by formally maximizing the SIR, we can derive the analytic solution for all arbitrary cases. We will closely follow the derivation given in both [7] and [30].

The general non-adaptive conventional narrowband array is shown in Fig. 8.5.

Figure 8.5 shows one desired signal arriving from the angle θ_0 and N interferers arriving from angles $\theta_1, \dots, \theta_N$. The signal and the interferers are received by an array of M elements with M potential weights. Each received signal at element m also includes additive Gaussian noise. Time is represented by the k th time sample. Thus, the weighted

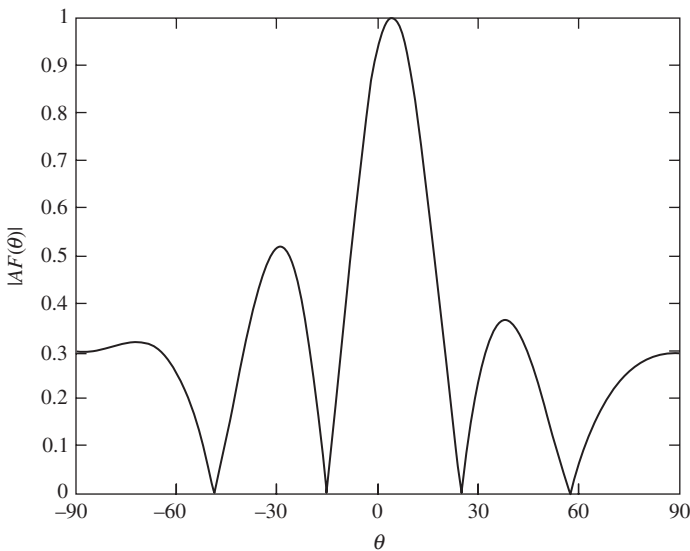


Figure 8.4 Array pattern with approximate nulls at -15° and 25° .

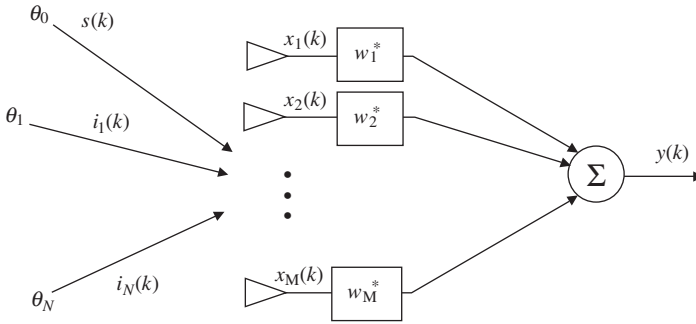


Figure 8.5 Traditional narrowband array.

array output y can be given in the following form:

$$y(k) = \bar{w}^H \cdot \bar{x}(k) \quad (8.9)$$

where

$$\begin{aligned} \bar{x}(k) &= \bar{a}_0 s(k) + [\bar{a}_1 \quad \bar{a}_2 \quad \dots \quad \bar{a}_N] \cdot \begin{bmatrix} i_1(k) \\ i_2(k) \\ \vdots \\ i_N(k) \end{bmatrix} + \bar{n}(k) \\ &= \bar{x}_s(k) + \bar{x}_i(k) + \bar{n}(k) \end{aligned} \quad (8.10)$$

with

$\bar{w} = [w_1 \ w_2 \ \dots \ w_M]^T$ = array weights

$\bar{x}_s(k)$ = desired signal vector

$\bar{x}_i(k)$ = interfering signals vector

$\bar{n}(k)$ = zero mean Gaussian noise for each channel

$\bar{a}_i$ = M -element array steering vector for the θ_i direction of arrival

We may rewrite Eq.(8.9) using the expanded notation in Eq. (8.10)

$$y(k) = \bar{w}^H \cdot [\bar{x}_s(k) + \bar{x}_i(k) + \bar{n}(k)] = \bar{w}^H \cdot [\bar{x}_s(k) + \bar{u}(k)] \quad (8.11)$$

where

$$\bar{u}(k) = \bar{x}_i(k) + \bar{n}(k) = \text{undesired signal}$$

It is initially assumed that all arriving signals are monochromatic and the total number of arriving signals $N + 1 \leq M$. It is understood that the arriving signals are time varying and thus our calculations

are based upon k -time snapshots of the incoming signal. Obviously, if the emitters are moving, the matrix of steering vectors is changing with time and the corresponding arrival angles are changing. Unless otherwise stated, the time dependence will be suppressed in Eq. (8.9) through (8.11).

We can calculate the array correlation matrices for both the desired signal ($\bar{R}_{ss}$) and the undesired signals ($\bar{R}_{uu}$). The literature often calls these matrices the *array covariance* matrices. However, the covariance matrix is a mean removed correlation matrix. Since we do not generally know the statistical mean of the system noise or the front end detector output, it is best to label all $\bar{R}$ matrices as correlation matrices. If the process is ergodic and the time average is utilized, the correlation matrices can be defined with the time average notation as $\hat{R}_{ss}$ and $\hat{R}_{uu}$.

The weighted array output power for the desired signal is given by

$$\sigma_s^2 = E \left[|\bar{w}^H \cdot \bar{x}_s|^2 \right] = \bar{w}^H \cdot \bar{R}_{ss} \cdot \bar{w} \quad (8.12)$$

where

$$\bar{R}_{ss} = E [\bar{x}_s \bar{x}_s^H] = \text{signal correlation matrix}$$

The weighted array output power for the undesired signals is given by

$$\sigma_u^2 = E [|\bar{w}^H \cdot \bar{u}|^2] = \bar{w}^H \cdot \bar{R}_{uu} \cdot \bar{w} \quad (8.13)$$

where it can be shown that

$$\bar{R}_{uu} = \bar{R}_{ii} + \bar{R}_{nn} \quad (8.14)$$

with

$$\bar{R}_{ii} = \text{correlation matrix for interferers}$$

$$\bar{R}_{nn} = \text{correlation matrix for noise}$$

The (SIR) is defined as the ratio of the desired signal power divided by the undesired signal power.

$$SIR = \frac{\sigma_s^2}{\sigma_u^2} = \frac{\bar{w}^H \cdot \bar{R}_{ss} \cdot \bar{w}}{\bar{w}^H \cdot \bar{R}_{uu} \cdot \bar{w}} \quad (8.15)$$

The SIR can be maximized in Eq. (8.15) by taking the derivative with respect to $\bar{w}$ and setting the result equal to zero. This optimization procedure is outlined in Harrington [31]. Rearranging terms, we can derive the following relationship

$$\bar{R}_{ss} \cdot \bar{w} = SIR \cdot \bar{R}_{uu} \cdot \bar{w} \quad (8.16)$$

or

$$\bar{R}_{uu}^{-1} \bar{R}_{ss} \cdot \bar{w} = SIR \cdot \bar{w} \quad (8.17)$$

Equation (8.17) is an eigenvector equation with SIR being the eigenvalues. The maximum SIR ($SIR_{\max}$) is equal to the largest eigenvalue $\lambda_{\max}$ for the Hermitian matrix $\bar{R}_{uu}^{-1}\bar{R}_{ss}$. The eigenvector associated with the largest eigenvalue is the optimum weight vector $\bar{w}_{\text{opt}}$. Thus

$$\bar{R}_{uu}^{-1}\bar{R}_{ss} \cdot \bar{w}_{SIR} = \lambda_{\max} \cdot \bar{w}_{\text{opt}} = SIR_{\max} \cdot \bar{w}_{SIR} \quad (8.18)$$

Since the correlation matrix is defined as $\bar{R}_{ss} = E[|s|^2]\bar{a}_0 \cdot \bar{a}_0^H$, we can pose the weight vector in terms of the optimum Wiener solution.

$$\bar{w}_{SIR} = \beta \cdot \bar{R}_{uu}^{-1} \cdot \bar{a}_0 \quad (8.19)$$

where

$$\beta = \frac{E[|s|^2]}{SIR_{\max}} \bar{a}_0^H \cdot \bar{w}_{SIR} \quad (8.20)$$

Although Eq. (8.19) casts the weight vector in the optimum Wiener solution form, the weight vector is already known as the eigenvector found in Eq. (8.18).

Example 8.2 The $M = 3$ -element array with spacing $d = .5\lambda$ has a noise variance $\sigma_n^2 = .001$, a desired received signal arriving at $\theta_0 = 30^\circ$, and two interferers arriving at angles $\theta_1 = -30^\circ$ and $\theta_2 = 45^\circ$. Assume that the signal and interferer amplitudes are constant. Use MATLAB to calculate $SIR_{\max}$, the normalized weights from Eq. (8.18), and plot the resulting pattern.

Solution Based upon the incident angles of arrival for the desired signal and interferers along with the array vector $\bar{a}$, we can find the correlation matrices of the signal and undesired signals as

$$\bar{R}_{ss} = \begin{bmatrix} 1 & i & -1 \\ -i & 1 & i \\ -1 & -i & 1 \end{bmatrix}$$

$$\bar{R}_{uu} = \begin{bmatrix} 2.001 & -.61 - .20i & -1.27 - .96i \\ -.61 + .20i & 2.001 & -.61 - .20i \\ -1.27 + .96i & -.61 + .20i & 2.001 \end{bmatrix}$$

The largest eigenvalue for Eq. (8.18) is given in MATLAB as

$$SIR_{\max} = \lambda_{\max} = 679$$

The array weights are arbitrarily normalized by the center weight value. Thus

$$\bar{w}_{SIR} = \begin{bmatrix} 1.48 + .5i \\ 1 \\ 1.48 - .5i \end{bmatrix}$$

The derived pattern is shown plotted in Fig. 8.6.

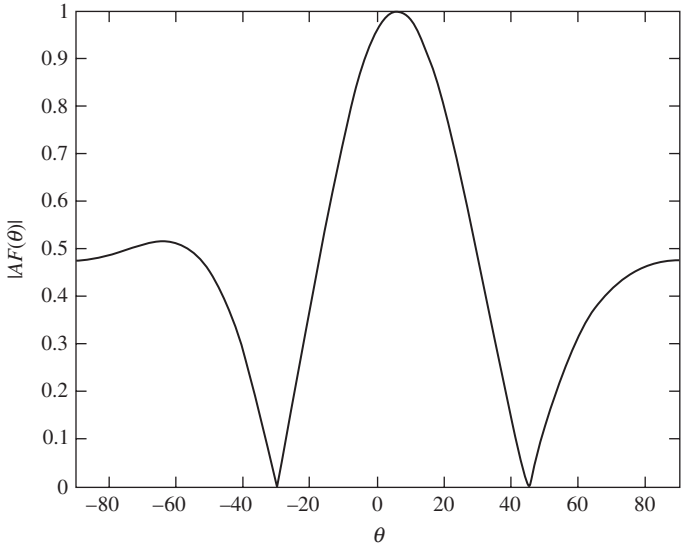


Figure 8.6 Maximum SIR pattern.

8.3.2 Minimum mean-square error

An alternative means for optimizing the array weights is found by minimizing the MSE. Figure 8.5 must be modified in such a way as to minimize the error while iterating the array weights. The modified array configuration is shown in Fig. 8.7.

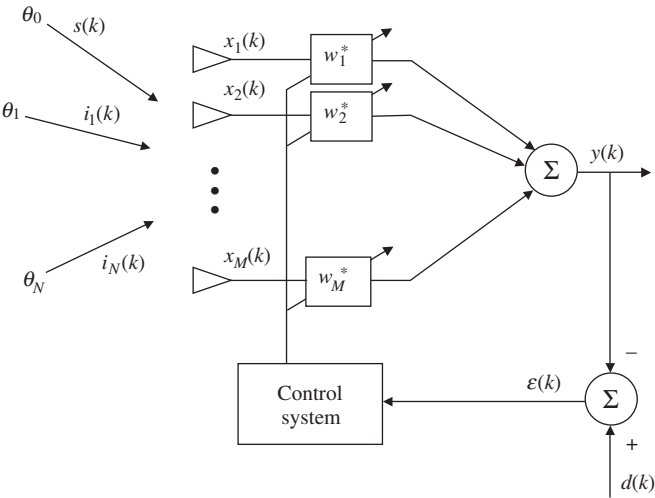


Figure 8.7 MSE adaptive system.

The signal $d(k)$ is the reference signal. Preferably the reference signal is either identical to the desired signal $s(k)$ or it is highly correlated with $s(k)$ and uncorrelated with the interfering signals $i_n(k)$. If $s(k)$ is not distinctly different from the interfering signals, the minimum mean-square technique will not work properly. The signal $\varepsilon(k)$ is the error signal such that

$$\varepsilon(k) = d(k) - \bar{w}^H \bar{x}(k) \quad (8.21)$$

Through some simple algebra, it can be shown that the MSE is given by

$$|\varepsilon(k)|^2 = |d(k)|^2 - 2d(k)\bar{w}^H \bar{x}(k) + \bar{w}^H \bar{x}(k) \bar{x}^H(k) \bar{w} \quad (8.22)$$

For the purposes of simplification, we will suppress the time dependence notation k . Taking the expected value of both sides and simplifying the expression we get

$$E[|\varepsilon|^2] = E[|d|^2] - 2\bar{w}^H \bar{r} + \bar{w}^H \bar{R}_{xx} \bar{w} \quad (8.23)$$

where the following correlations are defined:

$$\bar{r} = E[d^* \cdot \bar{x}] = E[d^* \cdot (\bar{x}_s + \bar{x}_i + \bar{n})] \quad (8.24)$$

$$\bar{R}_{xx} = E[\bar{x} \bar{x}^H] = \bar{R}_{ss} + \bar{R}_{uu} \quad (8.25)$$

$$\bar{R}_{ss} = E[\bar{x}_s \bar{x}_s^H] \quad (8.26)$$

$$\bar{R}_{uu} = \bar{R}_{ii} + \bar{R}_{nn} \quad (8.27)$$

The expression in Eq. (8.23) is a quadratic function of the weight vector $\bar{w}$. This function is sometimes called the *performance surface* or *cost function* and forms a quadratic surface in M -dimensional space. Since the optimum weights provide the minimum MSE, the extremum is the minimum of this function. A trivial example is given for a two-element array that produces a two-dimensional surface as depicted in Fig. 8.8. As the desired angle-of-arrival (AOA) changes with time, the quadratic surface minimum changes with time in the weight plane.

In general, for an arbitrary number of weights, we can find the minimum value by taking the gradient of the MSE with respect to the weight vectors and equating it to zero. Thus the Wiener-Hopf equation is given as

$$\nabla_{\bar{w}}(E[|\varepsilon|^2]) = 2\bar{R}_{xx} \bar{w} - 2\bar{r} = 0 \quad (8.28)$$

Simple algebra can be applied to Eq. (8.28) to yield the optimum Wiener solution given as

$$\bar{w}_{MSE} = \bar{R}_{xx}^{-1} \bar{r} \quad (8.29)$$

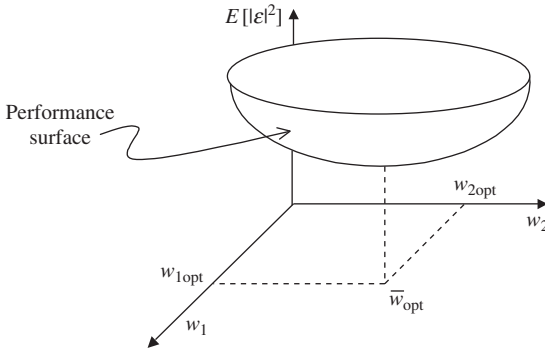


Figure 8.8 Quadratic surface for MSE.

If we allow the reference signal d to be equal to the desired signal s , and if s is uncorrelated with all interferers, we may simplify the correlation $\bar{r}$. Using Eqs. (8.10) and (8.24) the simplified correlation $\bar{r}$ is

$$\bar{r} = E[s * \bar{x}] = S \cdot \bar{a}_0 \quad (8.30)$$

where

$$S = E[|s|^2]$$

The optimum weights can then be identified as

$$\bar{w}_{MSE} = S \bar{R}_{xx}^{-1} \bar{a}_0 \quad (8.31)$$

Example 8.3 The $M = 5$ -element array with spacing $d = .5\lambda$ has a received signal energy $S = 1$ arriving at $\theta_0 = 20^\circ$, and two interferers arriving at angles $\theta_1 = -20^\circ$ and $\theta_2 = 40^\circ$, with noise variance $\sigma_n^2 = .001$. Use MATLAB to calculate the optimum weights and plot the resulting pattern.

Solution MATLAB code sa_ex8_3.m is used. The array vectors for the desired signal and the two interferers are given by

$$\bar{a}_0 = \begin{bmatrix} -.55 - .84i \\ .48 - .88i \\ 1.0 \\ .48 + .88i \\ -.55 + .84i \end{bmatrix} \quad \bar{a}_1 = \begin{bmatrix} -.55 + .84i \\ .48 + .88i \\ 1.0 \\ .48 - .88i \\ -.55 - .84i \end{bmatrix} \quad \bar{a}_2 = \begin{bmatrix} -.62 + .78i \\ -.43 - .90i \\ 1.0 \\ -.43 + .90i \\ -.62 - .78i \end{bmatrix}$$

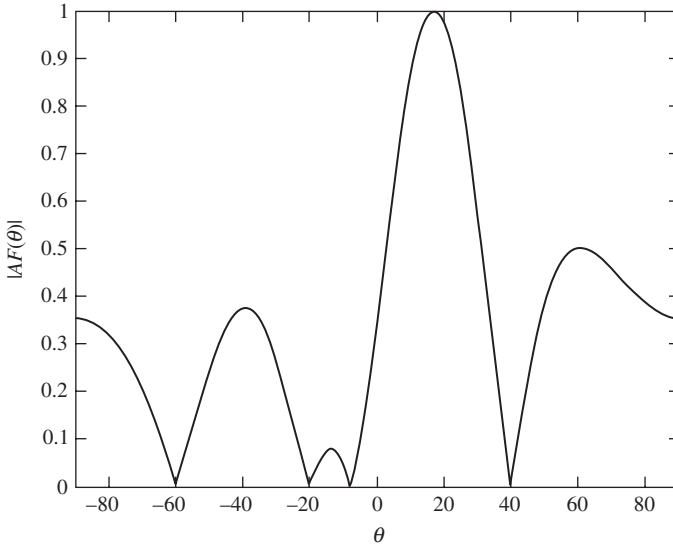


Figure 8.9 Minimum MSE pattern for 5-element array.

The correlation matrix for the array is given by

$$\bar{R}_{xx} = \begin{bmatrix} 3.0 & .51 - .90i & -1.71 + .78i & -1.01 + .22i & -1.02 - .97i \\ .51 + .90i & 3.0 & .51 - .90i & -1.71 + .78i & -1.01 + .22i \\ -1.71 - .78i & .51 + .90i & 3.0 & .51 - .90i & -1.71 + .78i \\ -1.01 - .22i & -1.71 - .78i & .51 + .90i & 3.0 & .51 - .90i \\ -1.02 + .97i & -1.01 - .22i & -1.71 - .78i & .51 + .90i & 3.0 \end{bmatrix}$$

The weights can now be calculated by using Eq. (8.25) to get

$$\bar{w}_{MSE} = \begin{bmatrix} -.11 - .21i \\ .18 - .08i \\ .21 \\ .18 + .08i \\ -.11 + .21i \end{bmatrix}$$

Applying the weights to the array vector, we can plot the resulting minimum MSE pattern Fig. 8.9.

8.3.3 Maximum likelihood

The maximum likelihood (ML) method is predicated on the assumption that we have an unknown desired signal $\bar{x}_s$ and that the unwanted signal $\bar{n}$ has a zero mean Gaussian distribution. The goal of this method is to define a likelihood function which can give us an estimate on the

desired signal. Details of the ML approach are found in excellent treatments by Van Trees [32, 33]. This fundamental solution falls under the general category of estimation theory. Referring to Fig. 8.10 it should be noted that no feedback is given to the antenna elements. The input signal vector is given by

$$\bar{x} = \bar{a}_0 s + \bar{n} = \bar{x}_s + \bar{n} \quad (8.32)$$

The overall distribution is assumed to be Gaussian but the mean is controlled by the desired signal $\bar{x}_s$. The probability density function can be described as the joint probability density $p(\bar{x}|\bar{x}_s)$. This density can be viewed as the likelihood function (Haykin [34] and Monzingo [30]) that can be used to estimate the parameter $\bar{x}_s$. The *probability density* can be described as

$$p(\bar{x}|\bar{x}_s) = \frac{1}{\sqrt{2\pi\sigma_n^2}} e^{-((\bar{x}-\bar{a}_0 s)^H \bar{R}_{nn}^{-1} (\bar{x}-\bar{a}_0 s))} \quad (8.33)$$

where σ_n = noise standard deviation

$\bar{R}_{nn} = \sigma_n^2 \bar{I}$ = noise correlation matrix

Since the parameter of interest is in the exponent, it is easier to work with the negative of the logarithm of the density function. We shall call this the *log-likelihood function*. Thus, we can define the log-likelihood function as

$$L[\bar{x}] = -\ln [p(\bar{x}|\bar{x}_s)] = C(\bar{x} - \bar{a}_0 s)^H \bar{R}_{nn}^{-1} (\bar{x} - \bar{a}_0 s) \quad (8.34)$$

where C = constant

$$\bar{R}_{nn} = E[\bar{n}\bar{n}^H]$$

Let us define our estimate of the desired signal, called $\hat{s}$, that maximizes the log-likelihood function. The maximum of $L[\bar{x}]$ is found by taking the partial derivative with respect to s and setting it equal to zero.

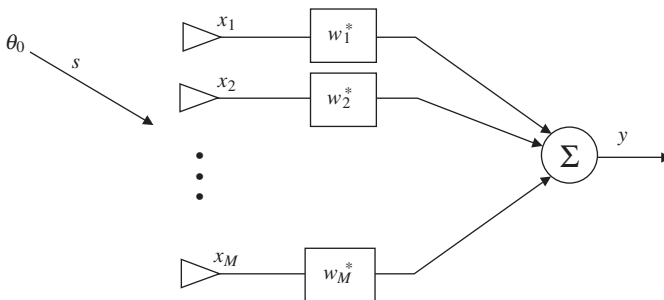


Figure 8.10 Traditional array.

Thus

$$\frac{\partial L[\bar{x}]}{\partial s} = 0 = -2\bar{a}_0^H \bar{R}_{nn}^{-1} \bar{x} + 2\hat{s}\bar{a}_0^H \bar{R}_{nn}^{-1} \bar{a}_0 \quad (8.35)$$

Solving for $\hat{s}$ we get

$$\hat{s} = \frac{\bar{a}_0^H \bar{R}_{nn}^{-1}}{\bar{a}_0^H \bar{R}_{nn}^{-1} \bar{a}_0} \bar{x} = \bar{w}_{ML}^H \bar{x} \quad (8.36)$$

Thus

$$\bar{w}_{ML} = \frac{\bar{R}_{nn}^{-1} \bar{a}_0}{\bar{a}_0^H \bar{R}_{nn}^{-1} \bar{a}_0} \quad (8.37)$$

Example 8.4 The $M = 5$ -element array with spacing $d = .5\lambda$ has a received signal arriving at the angle $\theta_0 = 30^\circ$, with noise variance $\sigma_n^2 = .001$. Use MATLAB to calculate the optimum weights and plot the resulting pattern.

Solution Since we are assuming that the noise is zero-mean Gaussian noise, the noise correlation matrix is the identity matrix given as

$$\bar{R}_{nn} = \sigma_n^2 \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

The array steering vector is given by

$$\bar{a}_0 = \begin{bmatrix} -1 \\ -1i \\ 1 \\ 1i \\ -1 \end{bmatrix}$$

The calculated array weights are given as

$$\bar{w}_{ML} = \frac{\bar{R}_{nn}^{-1} \bar{a}_0}{\bar{a}_0^H \bar{R}_{nn}^{-1} \bar{a}_0} = .2\bar{a}_0$$

The subsequent array factor plot is shown in Fig 8.11.

8.3.4 Minimum variance

The *minimum variance solution* [7] is sometimes called the minimum variance distortionless response (MVDR) [34] or the *minimum noise variance performance measure* [30]. The term *distortionless* is applied when it is desired that the received signal is undistorted after the application of the array weights. The goal of the MV method is to minimize

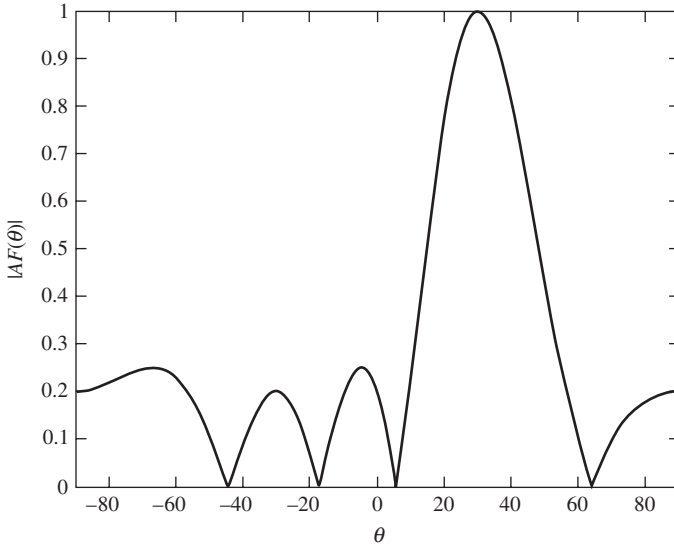


Figure 8.11 Maximum likelihood pattern for five element array.

the array output noise variance. It is assumed that the desired and unwanted signals have zero mean. We may again use the array configuration of Fig. 8.10. The weighted array output is given by

$$y = \bar{w}^H \bar{x} = \bar{w}^H \bar{a}_0 s + \bar{w}^H \bar{u} \quad (8.38)$$

In order to ensure a distortionless response, we must also add the constraint that

$$\bar{w}^H \bar{a}_0 = 1 \quad (8.39)$$

Applying the constraint to Eq. (8.38), the array output is given as

$$y = s + \bar{w}^H \bar{u} \quad (8.40)$$

In addition, if the unwanted signal has zero mean, the expected value of the array output is given by

$$E[y] = s \quad (8.41)$$

We may now calculate the variance of y such that

$$\begin{aligned} \sigma_{MV}^2 &= E[|\bar{w}^H \bar{x}|^2] = E[|s + \bar{w}^H \bar{u}|^2] \\ &= \bar{w}^H \bar{R}_{uu} \bar{w} \end{aligned} \quad (8.42)$$

where

$$\bar{R}_{uu} = \bar{R}_{ii} + \bar{R}_{nn}$$

We can minimize this variance by using the method of Lagrange [35]. Since all the array weights are interdependent, we can incorporate the constraint in Eq. (8.39) to define a *modified performance criterion*, or *cost function*, that is a linear combination of the variance and the constraint such that

$$\begin{aligned} J(\bar{w}) &= \frac{\sigma_{MV}^2}{2} + \lambda(1 - \bar{w}^H \bar{a}_0) \\ &= \frac{\bar{w}^H \bar{R}_{uu} \bar{w}}{2} + \lambda(1 - \bar{w}^H \bar{a}_0) \end{aligned} \quad (8.43)$$

where λ is the Lagrange multiplier and $J(\bar{w})$ is the cost function.

The cost function is a quadratic function and can be minimized by setting the gradient equal to zero. Thus

$$\nabla_{\bar{w}} J(\bar{w}) = \bar{R}_{uu} \bar{w}_{MV} - \lambda \bar{a}_0 = 0 \quad (8.44)$$

Solving for the weights, we conclude

$$\bar{w}_{MV} = \lambda \bar{R}_{uu}^{-1} \bar{a}_0 \quad (8.45)$$

In order to solve for the Lagrange multiplier (λ) we can substitute Eq. (8.39) into Eq. (8.45). Thus

$$\lambda = \frac{1}{\bar{a}_0^H \bar{R}_{uu}^{-1} \bar{a}_0} \quad (8.46)$$

Upon substituting Eq. (8.46) into Eq. (8.45), we arrive at the minimum variance optimum weights

$$\bar{w}_{MV} = \frac{\bar{R}_{uu}^{-1} \bar{a}_0}{\bar{a}_0^H \bar{R}_{uu}^{-1} \bar{a}_0} \quad (8.47)$$

It should be noted that the minimum variance solution is identical in form to the ML solution. The only difference is that the ML approach requires that all unwanted signals combined are zero mean and have a Gaussian distribution. However, with the minimum variance approach, the unwanted signal can include interferers arriving at unwanted angles as well as the noise. Thus, the minimum variance solution is more general in its application.

Example 8.5 The $M = 5$ -element array with spacing $d = .5\lambda$ has a received signal arriving at the angle $\theta_0 = 30^\circ$, one interferer arriving at -10° , and noise with a variance $\sigma_n^2 = .001$. Use MATLAB to calculate the optimum weights and plot the resulting pattern.

Solution It is a simple matter to slightly modify the MATLAB code from Example 8.3 to include the interferer at -10° . The MATLAB code used is

sa_ex8_5.m. The unwanted signal correlation matrix is given by

$$\bar{R}_{uu} = \begin{bmatrix} 1.0 & .85 + .52i & .46 + .89i & -.07 + .99i & -.57 + .82i \\ .85 - .52i & 1.0 & .85 + .52i & .46 + .89i & -.07 + .99i \\ .46 - .89i & .85 - .52i & 1.0 & .85 + .52i & .46 + .89i \\ -.07 - .99i & .46 - .89i & .85 - .52i & 1.0 & .85 + .52i \\ -.57 - .82i & -.07 - .99i & .46 - .89i & .85 - .52i & 1.0 \end{bmatrix}$$

The desired array vector is

$$\bar{a}(\theta_0) = \begin{bmatrix} -1 \\ -1i \\ 1 \\ 1i \\ -1 \end{bmatrix}$$

Using Eq. (8.40), the calculated minimum variance weights are given as

$$\bar{w} = \begin{bmatrix} -.19 + .04i \\ .03 - .19i \\ .25 \\ .03 + .19i \\ -.19 - .04i \end{bmatrix}$$

The subsequent array factor plot, with cancellation for interference, is shown in Fig 8.12. It should be noted that the minimum variance solution allows for canceling interferers and a null has been placed near -10° .

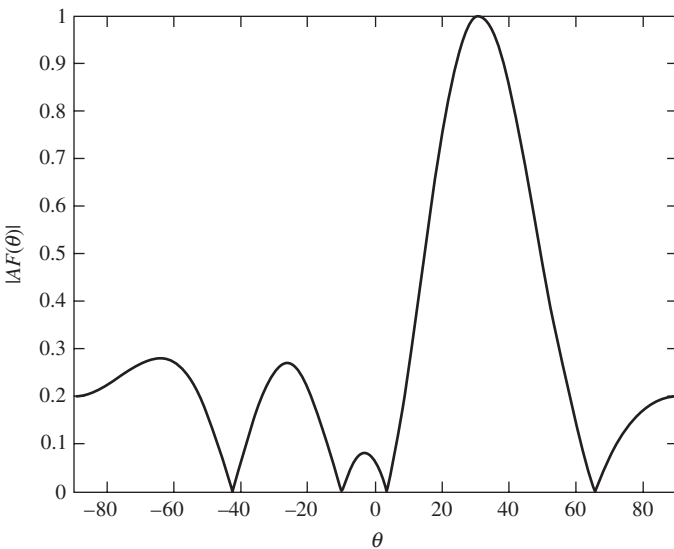


Figure 8.12 Minimum variance pattern for 5-element array.

8.4 Adaptive Beamforming

The fixed beamforming approaches, mentioned in Sec. 8.3, which included the maximum SIR, the ML method, and the MV method were assumed to apply to fixed arrival angle emitters. If the arrival angles don't change with time, the optimum array weights won't need to be adjusted. However, if the desired arrival angles change with time, it is necessary to devise an optimization scheme that operates *on-the-fly* so as to keep recalculating the optimum array weights. The receiver signal processing algorithm then must allow for the continuous adaptation to an ever-changing electromagnetic environment. The adaptive algorithm takes the fixed beamforming process one step further and allows for the calculation of continuously updated weights. The adaptation process must satisfy a specified optimization criterion. Several examples of popular optimization techniques include LMS, SMI, *recursive least squares* (RLS), the *constant modulus algorithm* (CMA), *conjugate gradient*, and waveform diverse algorithms. We will discuss and explain each of these techniques in the following sections.

8.4.1 Least mean squares

The least mean squares algorithm is a gradient based approach. Monzingo [30] gives an excellent fundamental treatment of this approach. Gradient based algorithms assume an established quadratic performance surface such as discussed in Sec. 8.3.2. When the performance surface is a quadratic function of the array weights, the performance surface $J(\bar{w})$ is in the shape of an elliptic paraboloid having one minimum. One of the best ways to establish the minimum is through the use of a gradient method. We can establish the performance surface (cost function) by again finding the MSE. The error, as indicated in Fig. 8.7, is

$$\varepsilon(k) = d(k) - \bar{w}^H(k)\bar{x}(k) \quad (8.48)$$

The squared error is given as

$$|\varepsilon(k)|^2 = |d(k) - \bar{w}^H(k)\bar{x}(k)|^2 \quad (8.49)$$

Momentarily, we will suppress the time dependence. As calculated in Sec. 8.3.2, the cost function is given as

$$J(\bar{w}) = D - 2\bar{w}^H \bar{r} + \bar{w}^H \bar{R}_{xx} \bar{w} \quad (8.50)$$

where

$$D = E[|d|^2]$$

We may employ the gradient method to locate the minimum of Eq. (8.50). Thus

$$\nabla_{\bar{w}} (J(\bar{w})) = 2\bar{R}_{xx}\bar{w} - 2\bar{r} \quad (8.51)$$

The minimum occurs when the gradient is zero. Thus, the solution for the weights is the optimum Wiener solution as given by

$$\bar{w}_{\text{opt}} = \bar{R}_{xx}^{-1}\bar{r} \quad (8.52)$$

The solution in Eq. (8.52) is predicated on our knowledge of all signal statistics and thus in our calculation of the correlation matrix.

In general, we do not know the signal statistics and thus must resort to estimating the array correlation matrix ($\bar{R}_{xx}$) and the signal correlation vector ($\bar{r}$) over a range of snapshots or for each instant in time. The instantaneous estimates of these values are given as

$$\hat{R}_{xx}(k) \approx \bar{x}(k)\bar{x}^H(k) \quad (8.53)$$

and

$$\hat{r}(k) \approx d^*(k)\bar{x}(k) \quad (8.54)$$

We can employ an iterative technique called the method of *steepest descent* to approximate the gradient of the cost function. The direction of steepest descent is in the opposite direction as the gradient vector. The method of steepest descent can be approximated in terms of the weights using the LMS method advocated by Widrow [16, 17]. The steepest descent iterative approximation is given as

$$\bar{w}(k+1) = \bar{w}(k) - \frac{1}{2}\mu\nabla_{\bar{w}} (J(\bar{w}(k))) \quad (8.55)$$

where, μ is the step-size parameter and $\nabla_{\bar{w}}$ is the gradient of the performance surface.

The gradient of the performance surface is given in Eq. (8.51). If we substitute the instantaneous correlation approximations, we have the LMS solution.

$$\begin{aligned} \bar{w}(k+1) &= \bar{w}(k) - \mu[\hat{R}_{xx}\bar{w} - \hat{r}] \\ &= \bar{w}(k) + \mu e^*(k)\bar{x}(k) \end{aligned} \quad (8.56)$$

where

$$e(k) = d(k) - \bar{w}^H(k)\bar{x}(k) = \text{error signal}$$

The convergence of the LMS algorithm in Eq. (8.56) is directly proportional to the *step-size parameter* μ . If the step-size is too small, the

convergence is slow and we will have the overdamped case. If the convergence is slower than the changing angles of arrival, it is possible that the adaptive array cannot acquire the signal of interest fast enough to track the changing signal. If the step-size is too large, the LMS algorithm will overshoot the optimum weights of interest. This is called the *underdamped case*. If attempted convergence is too fast, the weights will oscillate about the optimum weights but will not accurately track the solution desired. It is therefore imperative to choose a step-size in a range that insures convergence. It can be shown that stability is insured provided that the following condition is met [30].

$$0 \leq \mu \leq \frac{1}{2\lambda_{\max}} \quad (8.57)$$

where $\lambda_{\max}$ is the largest eigenvalue of $\hat{R}_{xx}$.

Since the correlation matrix is positive definite, all eigenvalues are positive. If all the interfering signals are noise and there is only one signal of interest, we can approximate the condition in Eq. (8.57) as

$$0 \leq \mu \leq \frac{1}{2 \text{ trace}[\hat{R}_{xx}]} \quad (8.58)$$

Example 8.6 An $M = 8$ -element array with spacing $d = .5\lambda$ has a received signal arriving at the angle $\theta_0 = 30^\circ$, an interferer at $\theta_1 = -60^\circ$. Use MATLAB to write an LMS routine to solve for the desired weights. Assume that the desired received signal vector is defined by $\bar{x}_s(k) = \bar{a}_0 s(k)$, where $s(k) = \cos(2\pi t(k)/T)$; with $T = 1$ ms and $t = (1:100)*T/100$. Assume the interfering signal vector is defined by $\bar{x}_i(k) = \bar{a}_1 i(k)$, where $i(k) = \text{randn}(1,100)$. Both signals are nearly orthogonal over the time interval T . Let the desired signal $d(k) = s(k)$.

Use the least mean square algorithm given in Eq. (8.56) to solve for the optimum array weights. Assume that the initial array weights are all zero. Allow for 100 iterations. Using MATLAB:

- Let step size $\mu = .02$.
- Calculate the eight array weights for 100 iterations.
- Plot the resulting weights magnitude vs. iteration number.
- Plot the desired signal $s(k)$ and the array output $y(k)$.
- Plot the mean square error $|e|^2$.
- Plot the array factor using the final weights calculated.

Solution The MATLAB code `sa_ex8_6.m` is used to solve the problem. The magnitude of the weights vs. iteration number is shown in Fig. 8.13. Figure 8.14 shows how the array output acquires and tracks the desired signal after about 60 iterations. Figure 8.15 shows the resulting mean square error which converges to near zero after 60 iterations. Figure 8.16 shows the final weighted array which has a peak at the desired direction of 30° and a null at the interfering direction of -60° .

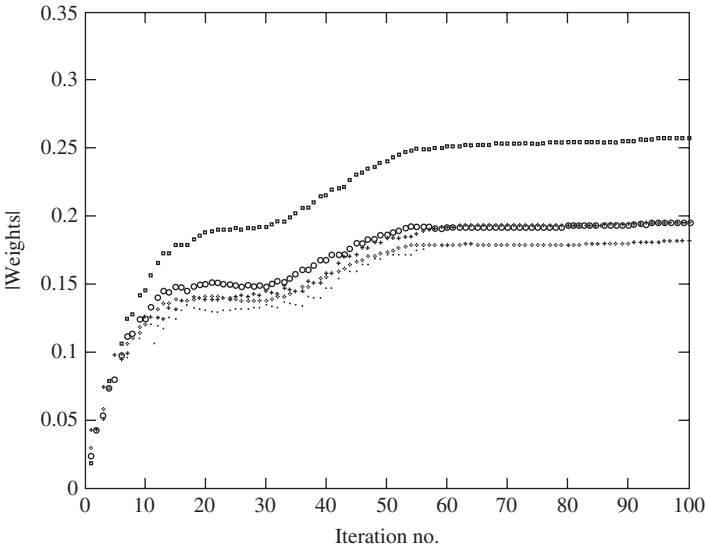


Figure 8.13 Magnitude of array weights.

8.4.2 Sample matrix inversion

One of the drawbacks of the LMS adaptive scheme is that the algorithm must go through many iterations before satisfactory convergence is achieved. If the signal characteristics are rapidly changing, the LMS

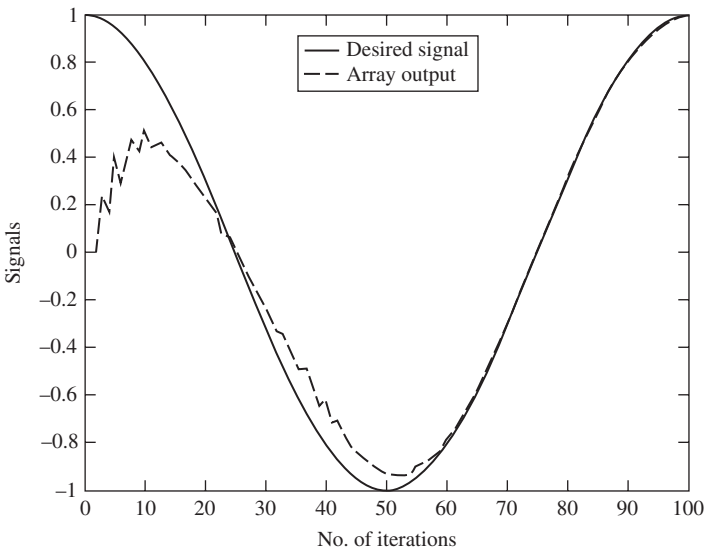


Figure 8.14 Acquisition and tracking of desired signal.

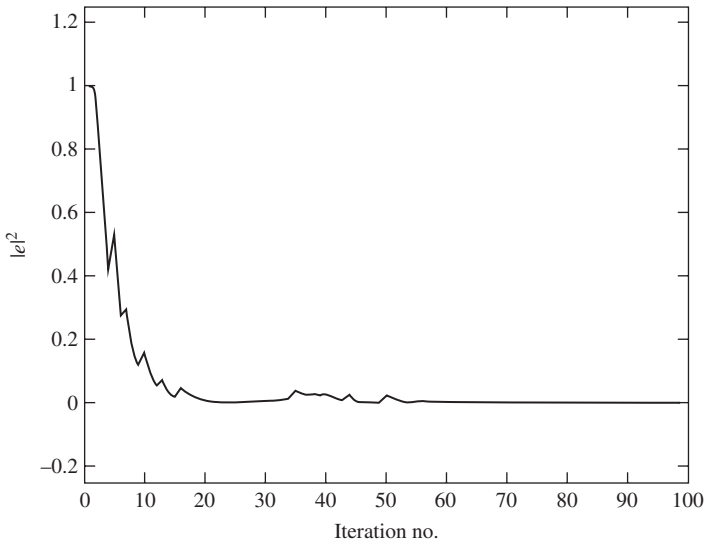


Figure 8.15 Mean square error.

adaptive algorithm may not allow tracking of the desired signal in a satisfactory manner. The rate of convergence of the weights is dictated by the eigenvalue spread of the array correlation matrix. In Example 8.6, the LMS algorithm did not converge until after 70 iterations. 70 iterations corresponded to more than half of the period of the waveform

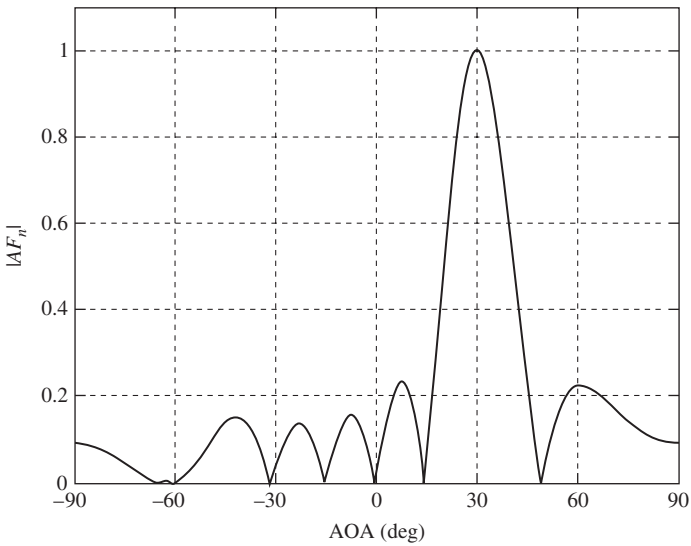


Figure 8.16 Weighted LMS array.

of interest. One possible approach to circumventing the relatively slow convergence of the LMS scheme is by use of SMI method [7, 23, 36]. This method is also alternatively known as *direct matrix inversion* (DMI) [30]. The *sample matrix* is a time average estimate of the array correlation, matrix using K -time samples. If the random process is ergodic in the correlation, the time average estimate will equal the actual correlation matrix.

Recalling the earlier discussion (Sec. 8.3.2) of the minimum MSE, the optimum array weights are given by the optimum Wiener solution as

$$\bar{w}_{\text{opt}} = \bar{R}_{xx}^{-1} \bar{r} \quad (8.59)$$

where

$$\begin{aligned} \bar{r} &= E [d^* \cdot \bar{x}] \\ \bar{R}_{xx} &= E [\bar{x} \bar{x}^H] \end{aligned}$$

As was shown with Eq. (7.32), we can estimate the correlation matrix by calculating the time average such that

$$\hat{R}_{xx} = \frac{1}{K} \sum_{k=1}^K \bar{x}(k) \bar{x}^H(k) \quad (8.60)$$

where K is the observation interval.

The correlation vector $\bar{r}$ can be estimated by

$$\hat{r} = \frac{1}{K} \sum_{k=1}^K d^*(k) \bar{x}(k) \quad (8.61)$$

Since we use a K -length block of data, this method is called a *block-adaptive approach*. We are thus adapting the weights block-by-block.

It is easy in MATLAB to calculate the array correlation matrix and the correlation vector by the following procedure. Define the matrix $\bar{X}_K(k)$ as the k th block of $\bar{x}$ vectors ranging over K -data snapshots. Thus

$$\bar{X}_K(k) = \begin{bmatrix} x_1(1+kK) & x_1(2+kK) & \cdots & x_1(K+kK) \\ x_2(1+kK) & x_2(2+kK) & & \vdots \\ \vdots & & \ddots & \\ x_M(1+kK) & \cdots & & x_M(K+kK) \end{bmatrix} \quad (8.62)$$

where k is the block number and K is the block length.

Thus, the estimate of the array correlation matrix is given by

$$\hat{R}_{xx}(k) = \frac{1}{K} \bar{X}_K(k) \bar{X}_K^H(k) \quad (8.63)$$

In addition, the desired signal vector can be define by

$$\bar{d}(k) = [d(1 + kK) \quad d(2 + kK) \quad \cdots \quad d(K + kK)] \quad (8.64)$$

Thus the estimate of the correlation vector is given by

$$\hat{r}(k) = \frac{1}{K} \bar{d}^*(k) \bar{X}_K^H(k) \quad (8.65)$$

The SMI weights can then be calculated for the k th block of length K as

$$\begin{aligned} \bar{w}_{SMI}(k) &= \bar{R}_{xx}^{-1}(k) \bar{r}(k) \\ &= [\bar{X}_K(k) \bar{X}_K^H(k)]^{-1} \bar{d}^*(k) \bar{X}_K(k) \end{aligned} \quad (8.66)$$

Example 8.7 Let us compare the SMI solution to the LMS solution of Example 8.5. An $M = 8$ -element array with spacing $d = .5\lambda$ has a received signal arriving at the angle $\theta_0 = 30^\circ$, an interferer at $\theta_1 = -60^\circ$. Use MATLAB to write an SMI routine to solve for the desired weights. Assume that the desired received signal vector is defined by $\bar{x}_s(k) = \bar{a}_0 s(k)$ where $s(k) = \cos(2\pi t(k)/T)$ with $T = 1$ ms. Let the block length be $K = 30$. Time is defined as $t = (1:K) \cdot T/K$. Assume the interfering signal vector is defined by $\bar{x}_i(k) = \bar{a}_1 i(k)$ where $i(k) = \text{randn}(1, K)$. Let the desired signal $d(k) = s(k)$. In order to keep the correlation matrix inverse from becoming singular, add noise to the system with variance $\sigma_n^2 = .01$.

Solution The MATLAB code `sa_ex8_7.m` is used to solve the problem. We may calculate the array input noise by use of the MATLAB command $n = \text{randn}(N, K) \cdot \text{sqrt}(\text{sig2})$; where sig2 = noise variance. After specifying the number of elements, the received signal angle, and the interfering angle, we calculate the received vector $\bar{x}(k)$. The correlation matrix is found using the simple MATLAB command $R_{xx} = X \cdot X^H / K$, the correlation vector is found using $r = X \cdot S^H / K$. The optimum Wiener weights are found and the resulting array pattern is seen in Fig. 8.17. The SMI pattern is similar to the LMS pattern and was generated with no iterations. The total number of snapshots K is less than the time to convergence for the LMS algorithm.

The SMI algorithm, although faster than the LMS algorithm, has several drawbacks. The correlation matrix may be ill conditioned resulting in errors or singularities when inverted. In addition, for large arrays, there is the challenge of inverting large matrices. To invert, the correlation matrix requires $N^3/2 + N^2$ complex multiplications [20]. The SMI update frequency will necessarily depend on signal frequencies and on channel fading conditions.

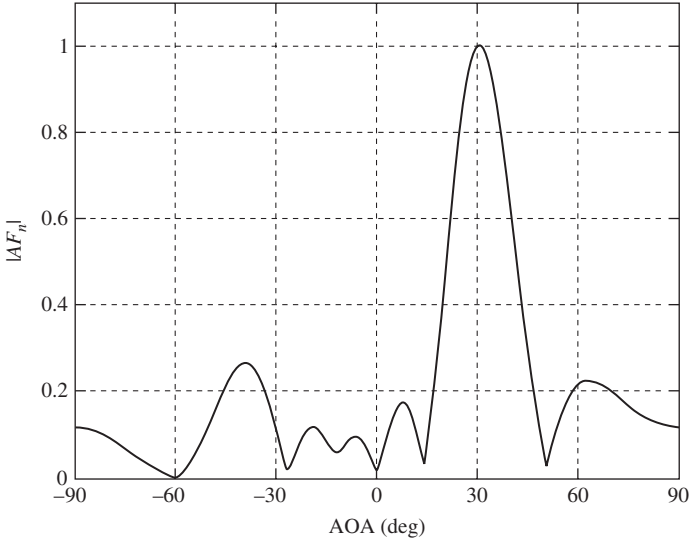


Figure 8.17 Weighted SMI array pattern.

8.4.3 Recursive least squares

As was mentioned in the previous section, the SMI technique has several drawbacks. Even though the SMI method is faster than the LMS algorithm, the computational burden and potential singularities can cause problems. However, we can recursively calculate the required correlation matrix and the required correlation vector. Recall that in Eqs. (8.60) and (8.61) the estimate of the correlation matrix and vector was taken as the sum of the terms divided by the block length K . When we calculate the weights in Eq. (8.66), the division by K is cancelled by the product $\bar{R}_{xx}^{-1}(k)\bar{r}(k)$. Thus, we can rewrite the correlation matrix and the correlation vector omitting K as

$$\hat{R}_{xx}(k) = \sum_{i=1}^k \bar{x}(i)\bar{x}^H(i) \tag{8.67}$$

$$\bar{r}(k) = \sum_{i=1}^k d^*(i)\bar{x}(i) \tag{8.68}$$

where k is the block length and last time sample k and $\hat{R}_{xx}(k), \bar{r}(k)$ is the correlation estimates ending at time sample k .

Both summations (Eqs. (8.67) and (8.68)) use rectangular windows, thus they equally consider all previous time samples. Since the signal sources can change or slowly move with time, we might want to deemphasize the earliest data samples and emphasize the most recent ones. This can be accomplished by modifying Eqs. (8.67) and (8.68) such that we forget the earliest time samples. This is called a *weighted estimate*.

Thus

$$\hat{R}_{xx}(k) = \sum_{i=1}^k \alpha^{k-i} \bar{x}(i) \bar{x}^H(i) \quad (8.69)$$

$$\hat{r}(k) = \sum_{i=1}^k \alpha^{k-i} d^*(i) \bar{x}(i) \quad (8.70)$$

where α is the *forgetting factor*.

The forgetting factor is also sometimes referred to as the *exponential weighting* factor [37]. α is a positive constant such that $0 \leq \alpha \leq 1$. When $\alpha = 1$, we restore the ordinary least squares algorithm. $\alpha = 1$ also indicates infinite memory. Let us break up the summation in Eqs. (8.69) and (8.70) into two terms: the summation for values up to $i = k-1$ and last term for $i = k$.

$$\begin{aligned} \hat{R}_{xx}(k) &= \alpha \sum_{i=1}^{k-1} \alpha^{k-1-i} \bar{x}(i) \bar{x}^H(i) + \bar{x}(k) \bar{x}^H(k) \\ &= \alpha \hat{R}_{xx}(k-1) + \bar{x}(k) \bar{x}^H(k) \end{aligned} \quad (8.71)$$

$$\begin{aligned} \hat{r}(k) &= \alpha \sum_{i=1}^{k-1} \alpha^{k-1-i} d^*(i) \bar{x}(i) + d^*(k) \bar{x}(k) \\ &= \alpha \hat{r}(k-1) + d^*(k) \bar{x}(k) \end{aligned} \quad (8.72)$$

Thus, future values for the array correlation estimate and the vector correlation estimate can be found using previous values.

Example 8.8 For an $M = 4$ -element array, $d = \lambda/2$, one signal arrives at 45° , and $S(k) = \cos(2\pi(k-1)/(K-1))$. Calculate the array correlation for a block of length $K = 200$ using the standard SMI algorithm and the recursion algorithm with $\alpha = 1$. Plot the trace of the SMI correlation matrix for K data points and the trace of the recursion correlation matrix vs. block length k where $1 < k < K$.

Solution Using MATLAB code `sa_ex8_8.m` we can construct the array steering vector for the angle of arrival of 45° . After multiplying the steering vector by the signal $S(k)$ we can then find the correlation matrix to start the recursion relationship in Eq. (8.71). After K iterations, we can superimpose the traces of both correlation matrices. This is shown in Fig. 8.18.

It can be seen that the recursion formula oscillates for different block lengths and that it matches the SMI solution when $k = K$. The recursion formula always gives a correlation matrix estimate for any block length k but only matches SMI when the forgetting factor is 1. The advantage of the recursion approach is that one need not calculate the correlation for an entire block of length K . Rather, each update only requires one a block of length 1 and the previous correlation matrix.

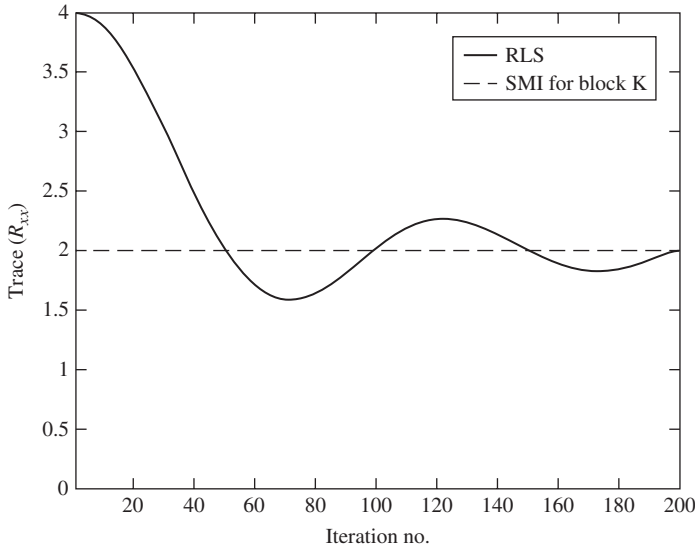


Figure 8.18 Trace of correlation matrix using SMI and RLS.

Not only can we recursively calculate the most recent correlation estimates, we can also use Eq. (8.71) to derive a recursion relationship for the inverse of the correlation matrix. The next steps follow the derivation in [37]. We can invoke the Sherman Morrison-Woodbury (SMW) theorem [38] to find the inverse of Eq. (8.71). Repeating the SMW theorem

$$(\bar{A} + \bar{z}\bar{z}^H)^{-1} = \bar{A}^{-1} - \frac{\bar{A}^{-1}\bar{z}\bar{z}^H\bar{A}^{-1}}{1 + \bar{z}^H\bar{A}^{-1}\bar{z}} \quad (8.73)$$

Applying Eq. (8.73) to Eq. (8.71), we have the following recursion formula:

$$\hat{R}_{xx}^{-1}(k) = \alpha^{-1}\hat{R}_{xx}^{-1}(k-1) - \frac{\alpha^{-2}\hat{R}_{xx}^{-1}(k-1)\bar{x}(k)\bar{x}^H(k)\hat{R}_{xx}^{-1}(k-1)}{1 + \alpha^{-1}\bar{x}^H(k)\hat{R}_{xx}^{-1}(k-1)\bar{x}(k)} \quad (8.74)$$

We can simplify Eq. (8.74) by defining the gain vector $\bar{g}(k)$

$$\bar{g}(k) = \frac{\alpha^{-1}\hat{R}_{xx}^{-1}(k-1)\bar{x}(k)}{1 + \alpha^{-1}\bar{x}^H(k)\hat{R}_{xx}^{-1}(k-1)\bar{x}(k)} \quad (8.75)$$

Thus

$$\hat{R}_{xx}^{-1}(k) = \alpha^{-1}\hat{R}_{xx}^{-1}(k-1) - \alpha^{-1}\bar{g}(k)\bar{x}^H(k)\hat{R}_{xx}^{-1}(k-1) \quad (8.76)$$

Equation (8.76) is known as the *Riccati equation* for the recursive least squares (RLS) method. We can rearrange Eq. (8.75) by multiplying the

denominator times both sides of the equation to yield

$$\bar{g}(k) = [\alpha^{-1} \hat{R}_{xx}^{-1}(k-1) - \alpha^{-1} \bar{g}(k) \bar{x}^H(k) \hat{R}_{xx}^{-1}(k-1)] \bar{x}(k) \quad (8.77)$$

It is clear that the term inside of the brackets of Eq. (8.77) is equal to Eq. (8.76). Thus

$$\bar{g}(k) = \hat{R}_{xx}^{-1}(k) \bar{x}(k) \quad (8.78)$$

Now we can derive a recursion relationship to update the weight vectors. The optimum Wiener solution is repeated in terms of the iteration number k and we can substitute Eq. (8.72) yielding

$$\begin{aligned} \bar{w}(k) &= \hat{R}_{xx}^{-1}(k) \hat{r}(k) \\ &= \alpha \hat{R}_{xx}^{-1}(k) \hat{r}(k-1) + \hat{R}_{xx}^{-1}(k) \bar{x}(k) d^*(k) \end{aligned} \quad (8.79)$$

We may now substitute Eq. (8.76) into the first correlation matrix inverse seen in Eq. (8.79).

$$\begin{aligned} \bar{w}(k) &= \hat{R}_{xx}^{-1}(k-1) \hat{r}(k-1) - \bar{g}(k) \bar{x}^H(k) \hat{R}_{xx}^{-1}(k-1) \hat{r}(k-1) \\ &\quad + \hat{R}_{xx}^{-1}(k) \bar{x}(k) d^*(k) \\ &= \bar{w}(k-1) - \bar{g}(k) \bar{x}^H(k) \bar{w}(k-1) + \hat{R}_{xx}^{-1}(k) \bar{x}(k) d^*(k) \end{aligned} \quad (8.80)$$

Finally we may substitute Eq. (8.78) into Eq. (8.80) to yield

$$\begin{aligned} \bar{w}(k) &= \bar{w}(k-1) - \bar{g}(k) \bar{x}^H(k) \bar{w}(k-1) + \bar{g}(k) d^*(k) \\ &= \bar{w}(k-1) + \bar{g}(k) [d^*(k) - \bar{x}^H(k) \bar{w}(k-1)] \end{aligned} \quad (8.81)$$

It should be noted that Eq. (8.81) is identical in form to Eq. (8.56).

Example 8.9 Use the RLS method to solve for the array weights and plot the resulting pattern. Let the array be an $M = 8$ -element array with spacing $d = .5\lambda$ with a received signal arriving at the angle $\theta_0 = 30^\circ$, an interferer at $\theta_1 = -60^\circ$. Use MATLAB to write an RLS routine to solve for the desired weights. Use Eqs. (8.71), (8.78), and (8.81). Assume that the desired received signal vector is defined by $\bar{x}_s(k) = \bar{a}_0 s(k)$ where $s(k) = \cos(2\pi t(k)/T)$; with $T = 1$ ms. Let there be $K = 50$ time samples such that $t = (0 : K-1) * T / (K-1)$. Assume that the interfering signal vector is defined by $\bar{x}_i(k) = \bar{a}_1 i(k)$ where $i(k) = \sin(\pi t(k)/T)$. Let the desired signal $d(k) = s(k)$. In order to keep the correlation matrix inverse from becoming singular, add noise to the system with variance $\sigma_n^2 = .01$. Begin with the assumption that all array weights are zero such that $\bar{w}(1) = [0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]^T$. Set the forgetting factor $\alpha = .9$.

Solution MATLAB code `sa_ex8_9.m` is used to solve for the array weights and to plot the resulting pattern shown in Fig. 8.19.

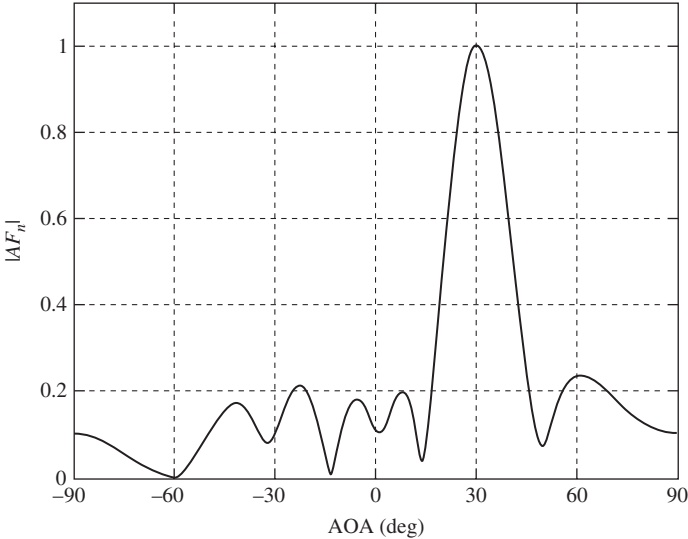


Figure 8.19 RLS array pattern.

The advantage of the RLS algorithm over SMI is that it is no longer necessary to invert a large correlation matrix. The recursive equations allow for easy updates of the inverse of the correlation matrix. The RLS algorithm also converges much more quickly than the LMS algorithm.

8.4.4 Constant modulus

Many adaptive beamforming algorithms are based on minimizing the error between a reference signal and the array output. The reference signal is typically a training sequence used to *train* the adaptive array or a desired signal based upon an a priori knowledge of nature of the arriving signals. In the case where a reference signal is not available one must resort to an assortment of optimization techniques that are *blind* to the exact content of the incoming signals.

Many wireless communication and radar signals are frequency-or phase-modulated signals. Some examples of phase and frequency-modulated signals are FM, PSK, FSK, QAM, and polyphase. This being the case, the amplitude of the signal should ideally be a constant. Thus the signal is said to have a constant magnitude or *modulus*. However, in fading channels, where multipath terms exist, the received signal is the composite of all multipath terms. Thus, the channel introduces an amplitude variation on the signal magnitude. Frequency selective channels by definition destroy the constant modulus property of the signal. If we know that the arriving signals of interest should have a constant modulus, we can devise algorithms that restore or *equalize* the amplitude of the original signal.

Dominique Godard [39] was the first to capitalize on the *constant modulus* (CM) property in order to create a family of blind equalization algorithms to be used in two-dimensional data communication systems. Specifically, Godard's algorithm applies to phase modulating waveforms. Godard used a cost function called a *dispersion function* of order p and, after minimization, the optimum weights are found. The Godard cost function is given by

$$J(k) = E [(|y(k)|^p - R_p)^q] \quad (8.82)$$

where p is the positive integer and q is the positive integer = 1.

Godard showed that the gradient of the cost function is zero when R_p is defined by

$$R_p = \frac{E [|s(k)|^{2p}]}{E [|s(k)|^p]} \quad (8.83)$$

where $s(k)$ is the zero-memory estimate of $y(k)$.

The resulting error signal is given by

$$e(k) = y(k)|y(k)|^{p-2}(R_p - |y(k)|^p) \quad (8.84)$$

This error signal can replace the traditional error signal in the LMS algorithm to yield

$$\bar{w}(k+1) = \bar{w}(k) + \mu e^*(k)\bar{x}(k) \quad (8.85)$$

The $p = 1$ case reduces the cost function to the form

$$J(k) = E [(|y(k)| - R_1)^2] \quad (8.86)$$

where

$$R_1 = \frac{E [|s(k)|^2]}{E [|s(k)|]} \quad (8.87)$$

If we scale the output estimate $s(k)$ to unity, we can write the error signal in Eq. (8.84) as

$$e(k) = \left(y(k) - \frac{y(k)}{|y(k)|} \right) \quad (8.88)$$

Thus the weight vector, in the $p = 1$ case, becomes

$$\bar{w}(k+1) = \bar{w}(k) + \mu \left(1 - \frac{1}{|y(k)|} \right) y^*(k)\bar{x}(k) \quad (8.89)$$

The $p = 2$ case reduces the cost function to the form

$$J(k) = E[(|y(k)|^2 - R_2)^2] \quad (8.90)$$

where

$$R_2 = \frac{E[|s(k)|^4]}{E[|s(k)|^2]} \quad (8.91)$$

If we scale the output estimate $s(k)$ to unity, we can write the error signal in Eq. (8.84) as

$$e(k) = y(k)(1 - |y(k)|^2) \quad (8.92)$$

Thus the weight vector, in the $p = 2$ case, becomes

$$\bar{w}(k+1) = \bar{w}(k) + \mu(1 - |y(k)|^2)y^*(k)\bar{x}(k) \quad (8.93)$$

The cases where $p = 1$ or 2 are referred to as *constant modulus algorithms* (CMA). The $p = 1$ case has been proven to converge much more rapidly than the $p = 2$ case [40]. A similar algorithm was developed by Treichler and Agee [41] and is identical to the Godard case for $p = 2$.

Example 8.10 Allow the same constant modulus signal to arrive at the receiver via a direct path and two additional multipaths and assume that the channel is frequency selective. Let the direct path arriving signal be defined as a 32-chip binary sequence where the chip values are ± 1 and are sampled four times per chip. The direct path signal arrives at 45° . The first multipath signal arrives at -30° but is 30 percent of the direct path in amplitude. The second multipath signal arrives at 0° but is 10 percent of the direct path in amplitude. Because of multipath, there will be slight time delays in the binary sequences causing dispersion. This time delay can be implemented by zero padding the multipath signals to effect a time delay. We will use the $P = 1$ CMA algorithm to define the optimum weights. Choose $\mu = .5$, $N = 8$ elements, and $d = \lambda/2$. Define the initial weights $\bar{w}(1)$ to be zero. Plot the resulting pattern.

Solution The three received waveforms are shown in Fig. 8.20. The last block indicates the combined received waveforms. It can be seen that the combined received signal has an amplitude variation due to channel dispersion.

The array output is defined as $y(k) = \bar{w}^H(k)\bar{x}(k)$. The recursion relationship for the array weights is given in Eq. (8.89). MATLAB code sa_ex8_10.m is used to solve for the weights. The resulting pattern plot is given in Fig. 8.21. It should be noted that the CMA algorithm suppresses the multipath but does not cancel it.

8.4.5 Least squares constant modulus

One severe disadvantage of the Godard CMA algorithm is the slow convergence time. The slow convergence limits the usefulness of the algorithm in dynamic environments where the signal must be captured

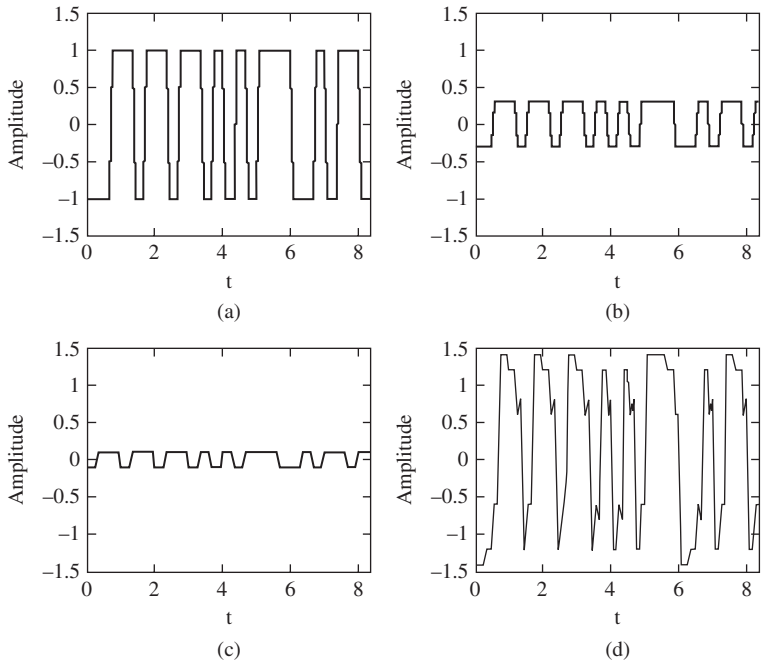


Figure 8.20 (a) Direct path, (b) path 2, (c) path 3, and (d) combined signals.

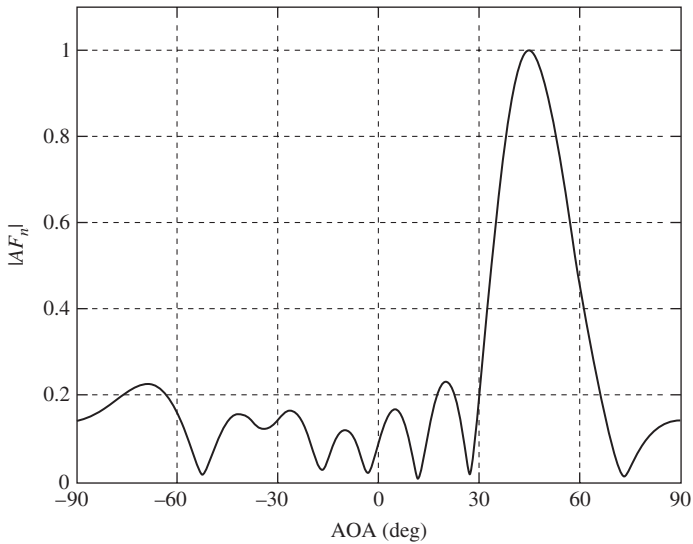


Figure 8.21 CMA pattern.

quickly. This also limits the usefulness of CMA when channel conditions are rapidly changing. The previous CMA method is based upon the method of steepest descent by taking the gradient of the cost function in Eq. (8.82). A faster algorithm was developed by Agee [42] using the method of nonlinear least-squares. The least-squares method is also known as the *Gauss method* based upon the work of Gauss in 1795 [43]. This method is referred to as the *LS-CMA algorithm* [44] and is also known as an *autoregressive estimator* based on a least squares minimization [45].

The following derivation is taken directly from [42] and [44]. In the method of least squares, one defines a cost function which is the weighted sum of error squares or the total error energy. The energies are the energies of a finite sample set K . The cost function is defined by

$$C(\bar{w}) = \sum_{k=1}^K |\phi_k(\bar{w})|^2 = \|\bar{\Phi}(\bar{w})\|_2^2 \quad (8.94)$$

where $\phi_k(\bar{w})$ = error at k th data sample

$$\bar{\Phi}(\bar{w}) = [\phi_1(\bar{w}) \quad \phi_1(\bar{w}) \quad \dots \quad \phi_K(\bar{w})]^T$$

K = number of data samples in one block

Equation (8.94) has a partial Taylor-series expansion with a sum-of-squares form given as

$$C(\bar{w} + \bar{\Delta}) \approx \|\bar{\Phi}(\bar{w}) + \bar{J}^H(\bar{w})\bar{\Delta}\|_2^2 \quad (8.95)$$

Where the complex Jacobian of $\bar{\Phi}(\bar{w})$ is defined as

$$\bar{J}(\bar{w}) = [\nabla\phi_1(\bar{w}) \quad \nabla\phi_2(\bar{w}) \quad \dots \quad \nabla\phi_K(\bar{w})] \quad (8.96)$$

and $\bar{\Delta}$ is the offset that updates weights.

We wish to find the offset $\bar{\Delta}$ which minimizes the sum-of-squared errors. Taking the gradient of Eq. (8.95) and setting it equal to zero, we can find the optimum offset vector to be defined as

$$\bar{\Delta} = -[\bar{J}(\bar{w})\bar{J}^H(\bar{w})]^{-1}\bar{J}(\bar{w})\bar{\Phi}(\bar{w}) \quad (8.97)$$

The new updated weight vector is then given by

$$\bar{w}(n+1) = \bar{w}(n) - [\bar{J}(\bar{w}(n))\bar{J}^H(\bar{w}(n))]^{-1}\bar{J}(\bar{w}(n))\bar{\Phi}(\bar{w}(n)) \quad (8.98)$$

The new weight vector is the previous weight vector adjusted by the offset $\bar{\Delta}$. The number n is the iteration number not to be confused with the time sample k .

Let us now apply the least squares method to the CMA using the 1-2 cost function [8].

$$C(\bar{w}) = \sum_{k=1}^K |\phi_k(\bar{w})|^2 = \sum_{k=1}^K ||y(k)| - 1|^2 \quad (8.99)$$

where $y(k) = \bar{w}^H \bar{x}(k)$ = array output at time k .
We may write ϕ_k as a vector such that

$$\bar{\phi}(\bar{w}) = \begin{bmatrix} |y(1)| - 1 \\ |y(2)| - 1 \\ \vdots \\ |y(K)| - 1 \end{bmatrix} \quad (8.100)$$

We may now define the Jacobian of the error vector $\bar{\phi}(\bar{w})$

$$\begin{aligned} \bar{J}(\bar{w}) &= [\nabla(\phi_1(\bar{w})) \quad \nabla(\phi_2(\bar{w})) \quad \dots \quad \nabla(\phi_K(\bar{w}))] \\ &= \begin{bmatrix} \bar{x}(1) \frac{y^*(1)}{|y(1)|} & \bar{x}(2) \frac{y^*(2)}{|y(2)|} & \dots & \bar{x}(K) \frac{y^*(K)}{|y(K)|} \end{bmatrix} \\ &= \bar{X} \bar{Y}_{CM} \end{aligned} \quad (8.101)$$

where

$$\bar{X} = [\bar{x}(1) \quad \bar{x}(2) \quad \dots \quad \bar{x}(K)] \quad (8.102)$$

and

$$\bar{Y}_{CM} = \begin{bmatrix} \frac{y^*(1)}{|y(1)|} & 0 & \dots & 0 \\ 0 & \frac{y^*(2)}{|y(2)|} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \frac{y^*(K)}{|y(K)|} \end{bmatrix} \quad (8.103)$$

Multiplying the Jacobian times its Hermitian transpose, we get

$$\bar{J}(\bar{w}) \bar{J}^H(\bar{w}) = \bar{X} \bar{Y}_{CM} \bar{Y}_{CM}^H \bar{X}^H = \bar{X} \bar{X}^H \quad (8.104)$$

The product of the Jacobian times the energy matrix is given by

$$\bar{J}(\bar{w}) \bar{\Phi}(\bar{w}) = \bar{X} \bar{Y}_{CM} \begin{bmatrix} |y(1)| - 1 \\ |y(2)| - 1 \\ \vdots \\ |y(K)| - 1 \end{bmatrix} = \bar{X} \begin{bmatrix} y^*(1) - \frac{y^*(1)}{|y(1)|} \\ y^*(2) - \frac{y^*(2)}{|y(2)|} \\ \vdots \\ y^*(K) - \frac{y^*(K)}{|y(K)|} \end{bmatrix} = \bar{X}(\bar{y} - \bar{r})^* \quad (8.105)$$

where

$$\bar{y} = [y(1) \quad y(2) \quad \dots \quad y(K)]^T \quad (8.106)$$

and

$$\bar{r} = \begin{bmatrix} \frac{y(1)}{|y(1)|} & \frac{y(2)}{|y(2)|} & \dots & \frac{y(K)}{|y(K)|} \end{bmatrix}^T = L(\bar{y}) \quad (8.107)$$

where $L(\bar{y})$ is a hard-limiter acting on $\bar{y}$.

Substituting Eqs. (8.103) and (8.105) into Eq. (8.98), we get

$$\begin{aligned} \bar{w}(n+1) &= \bar{w}(n) - [\bar{X}\bar{X}^H]^{-1} \bar{X}(\bar{y}(n) - \bar{r}(n))^* \\ &= \bar{w}(n) - [\bar{X}\bar{X}^H]^{-1} \bar{X}\bar{X}^H \bar{w}(n) + [\bar{X}\bar{X}^H]^{-1} \bar{X}\bar{r}^*(n) \\ &= [\bar{X}\bar{X}^H]^{-1} \bar{X}\bar{r}^*(n) \end{aligned} \quad (8.108)$$

where

$$\bar{r}^*(n) = \begin{bmatrix} \frac{\bar{w}^H(n)\bar{x}(1)}{|\bar{w}^H(n)\bar{x}(1)|} & \frac{\bar{w}^H(n)\bar{x}(2)}{|\bar{w}^H(n)\bar{x}(2)|} & \dots & \frac{\bar{w}^H(n)\bar{x}(K)}{|\bar{w}^H(n)\bar{x}(K)|} \end{bmatrix}^H \quad (8.109)$$

While only one block of data is used to implement the LS-CMA, the algorithm iterates through n values until convergence. Initial weights $\bar{w}(1)$ are chosen, the complex-limited output data vector $\bar{r}^*(1)$ is calculated, then the next weight $\bar{w}(2)$ is calculated, and the iteration continues until satisfactory convergence is satisfied. This is called the *static* LS-CMA algorithm because only one static block, of length K , is used for the iteration process. The LS-CMA algorithm bears a striking resemblance to the SMI algorithm in Eq. (8.66).

Example 8.11 Allow the same constant modulus signal to arrive at the receiver via a direct path and one additional multipath and assume that the channel is frequency selective. Let the direct path arriving signal be defined as a 32-bit binary chipping sequence where the chip values are ± 1 and are sampled four times per chip. Let the block length $K = 132$. The direct path signal arrives at 45° . A multipath signal arrives at -30° but is 30 percent of the direct path in amplitude. Because of multipath, there will be slight time delays in the binary sequences causing dispersion. Let the dispersion be implemented by zero padding the two signals. Zero pad the direct path by adding four zeros to the back of the signal. Zero pad the multipath signal with two zeros in front and two zeros in back. Use the LS-CMA algorithm to find the optimum weights. $N = 8$ elements and $d = \lambda/2$. Add zero mean Gaussian noise for each element in the array and let the noise variance be $\sigma_n^2 = .01$. Define the initial weights $\bar{w}(1)$ to be one. Iterate three times. Plot the resulting pattern.

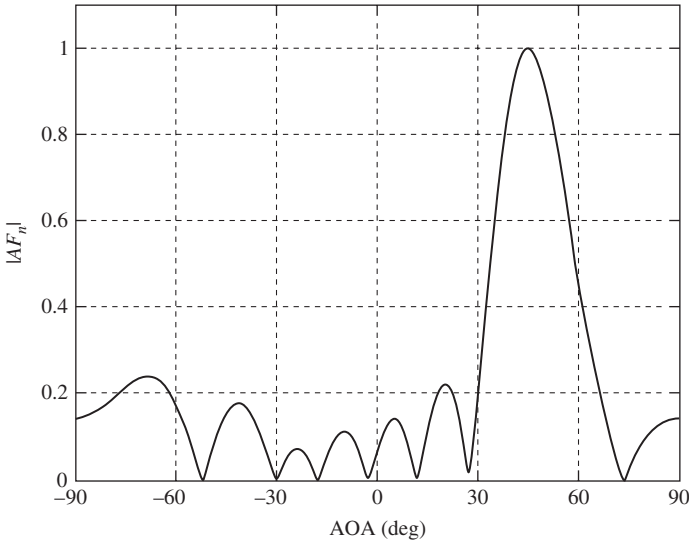


Figure 8.22 Static LS-CMA.

Solution The MATLAB code `sa_ex8_11.m` can be utilized for this purpose. The CM waveforms are generated identically to Example 8.10. The resulting pattern is shown in Fig. 8.22.

It should be noted that LS-CMA algorithm does a better job of nulling the multipath terms than does the CMA algorithm from the previous example.

The chief advantage of the static LS-CMA is that it can converge up to 100 times faster than the conventional CMA algorithm. In fact, in this example, the weights effectively converged after only a few iterations.

The static LS-CMA algorithm computed the weights simply based upon a fixed block of sampled data. In order to maintain up-to-date adaptation in a dynamic signal environment, it is better to update the data blocks for each iteration. Thus a dynamic LS-CMA algorithm is more appropriate. The *dynamic* LS-CMA is a modification of the previous static version. Let us define a dynamic block of data as the array output before applying weights. For the n th iteration, the n th block of length K is given as

$$\bar{X}(n) = [\bar{x}(1 + nK) \quad \bar{x}(2 + nK) \quad \cdots \quad \bar{x}(K + nK)] \quad (8.110)$$

The weighted array output, for the n th iteration is now defined as

$$\bar{y}(n) = [y(1 + nK) \quad y(2 + nK) \quad \cdots \quad y(K + nK)]^T = [\bar{w}^H(n) \bar{X}(n)]^T \quad (8.111)$$

The complex limited output data vector is given as

$$\bar{r}(n) = \left[\frac{y(1+nK)}{|y(1+nK)|} \quad \frac{y(2+nK)}{|y(2+nK)|} \quad \cdots \quad \frac{y(K+nK)}{|y(K+nK)|} \right]^T \quad (8.112)$$

Replacing Eq. (8.108) with the dynamic version we have

$$\bar{w}(n+1) = [\bar{X}(n)\bar{X}^H(n)]^{-1}\bar{X}(n)\bar{r}^*(n) \quad (8.113)$$

where

$$\bar{r}^*(n) = \left[\frac{\bar{w}^H(n)\bar{x}(1+nK)}{|\bar{w}^H(n)\bar{x}(1+nK)|} \quad \frac{\bar{w}^H(n)\bar{x}(2+nK)}{|\bar{w}^H(n)\bar{x}(2+nK)|} \quad \cdots \quad \frac{\bar{w}^H(n)\bar{x}(K+nK)}{|\bar{w}^H(n)\bar{x}(K+nK)|} \right]^H \quad (8.114)$$

We can further simplify Eq. (8.113) by defining the array correlation matrix and the correlation vector as

$$\hat{R}_{xx}(n) = \frac{\bar{X}(n)\bar{X}^H(n)}{K} \quad (8.115)$$

and

$$\hat{r}_{xr}(n) = \frac{\bar{X}(n)\bar{r}^*(n)}{K} \quad (8.116)$$

The dynamic LS-CMA is not defined as

$$\bar{w}(n+1) = \hat{R}_{xx}^{-1}(n)\hat{r}_{xr}(n) \quad (8.117)$$

Example 8.12 Repeat Example 8.11 using the dynamic LS-CMA. Define the block length $K = 22$ data points. Allow the block to update for every iteration n . $N = 8$ -elements and $d = \lambda/2$. Define the initial weights $\bar{w}(1)$ to be one. Solve for the weights after 6 iterations. Plot the resulting pattern.

Solution The MATLAB code `sa_ex8.12.m` is used for this purpose. The CM waveforms are generated identically to Example 8.10 except that the block of K -data points moves with the iteration number. The resulting pattern is shown in Fig. 8.23.

8.4.6 Conjugate gradient method

The problem with the steepest descent method has been the sensitivity of the convergence rates to the eigenvalue spread of the correlation matrix. Greater spreads result in slower convergences. The convergence rate can be accelerated by use of the *conjugate gradient method* (CGM). The goal of CGM is to iteratively search for the optimum solution by choosing conjugate (perpendicular) paths for each new iteration. Conjugacy in this context is intended to mean *orthogonal*. The method of

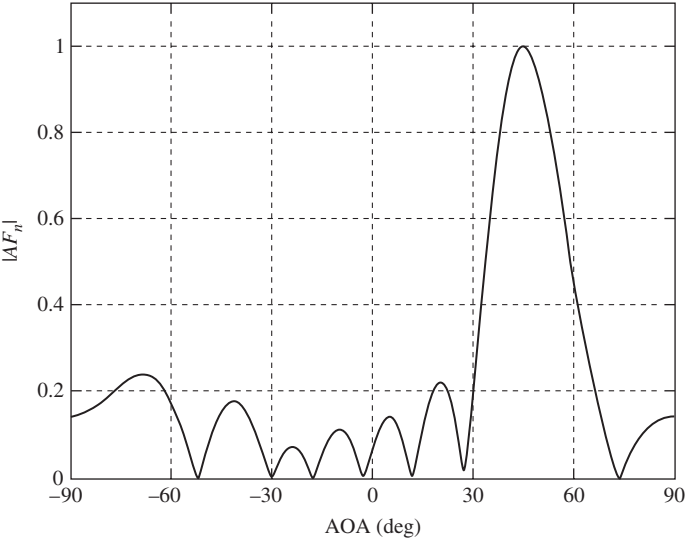


Figure 8.23 Dynamic LS-CMA.

CGM produces orthogonal search directions resulting in the fastest convergence. Figure 8.24 depicts a top view of a two-dimensional performance surface where the conjugate steps show convergence toward the optimum solution. Note that the path taken at iteration $n + 1$ is perpendicular to the path taken at the previous iteration n .

The CGM method has its early roots in the desire to solve a system of linear equations. One of the earliest references to CGM is found in a 1952 document by Hestenes and Stiefel [46]. Additionally, early work

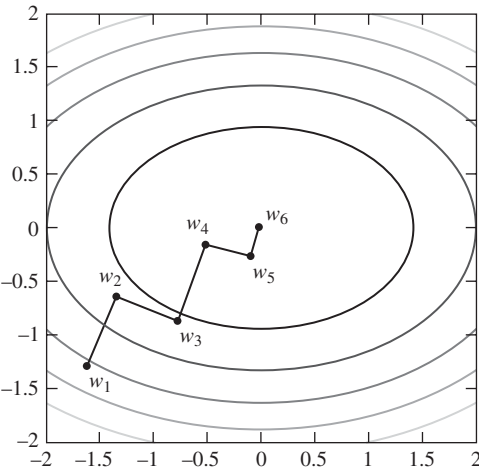


Figure 8.24 Convergence using conjugate directions.

was performed by Fletcher and Powell in 1963 [47] and Fletcher and Reeves in 1964 [48]. The CGM has also been called the *accelerated gradient approach* (AG) by Monzingo and Miller [30]. The gradient is accelerated by virtue of choosing conjugate directions. The CGM method was modified for use in predicting array weights by Choi [49, 50]. A survey of this method has been written by Godara [51] and a concise summary of the method can be found in Sadiku [52]. The following summary is taken from [51] and [52].

CGM is an iterative method whose goal is to minimize the quadratic cost function

$$J(\bar{w}) = \frac{1}{2} \bar{w}^H \bar{A} \bar{w} - \bar{d}^H \bar{w} \quad (8.118)$$

where

$$\bar{A} = \begin{bmatrix} x_1(1) & x_2(1) & \cdots & x_M(1) \\ x_1(2) & x_2(2) & \cdots & x_M(2) \\ \vdots & \vdots & \ddots & \vdots \\ x_1(K) & x_2(K) & \cdots & x_M(K) \end{bmatrix} \quad K \times M \text{ matrix of array snapshots}$$

K = number of snapshots

M = number of array elements

$\bar{w}$ = unknown weight vector

$\bar{d} = [d(1) \ d(2) \ \cdots \ d(K)]^T$ = desired signal vector of K snapshots

We may take the gradient of the cost function and set it to zero in order to find the minimum. It can be shown that

$$\nabla_{\bar{w}} J(\bar{w}) = \bar{A} \bar{w} - \bar{d} \quad (8.119)$$

We may employ the method of steepest descent in order to iterate to minimize Eq. (8.119). We wish to slide to the bottom of the quadratic cost function choosing the least number of iterations. We may start with an initial guess for the weights $\bar{w}(1)$ and find the residual $\bar{r}(1)$. The first residual value after at the first guess is given as

$$\bar{r}(1) = -J'(\bar{w}(1)) = \bar{d} - \bar{A} \bar{w}(1) \quad (8.120)$$

We can next choose a direction vector $\bar{D}$ which gives us the new conjugate direction to iterate toward the optimum weight. Thus

$$\bar{D}(1) = \bar{A}^H \bar{r}(1) \quad (8.121)$$

The general weight update equation is given by

$$\bar{w}(n+1) = \bar{w}(n) - \mu(n) \bar{D}(n) \quad (8.122)$$

Where the step size is determined by

$$\mu(n) = \frac{\bar{r}^H(n) \bar{A} \bar{A}^H \bar{r}(n)}{\bar{D}^H(n) \bar{A}^H \bar{A} \bar{D}(n)} \quad (8.123)$$

We may now update the residual and the direction vector. We can pre-multiply Eq. (8.122) by $-\bar{A}$ and add $\bar{d}$ to derive the updates for the residuals.

$$\bar{r}(n+1) = \bar{r}(n) + \mu(n) \bar{A} \bar{D}(n) \quad (8.124)$$

The direction vector update is given by

$$\bar{D}(n+1) = \bar{A}^H \bar{r}(n+1) - \alpha(n) \bar{D}(n) \quad (8.125)$$

We can use a linear search to determine $\alpha(n)$ which minimizes $J(\bar{w}(n))$. Thus

$$\alpha(n) = \frac{\bar{r}^H(n+1) \bar{A} \bar{A}^H \bar{r}(n+1)}{\bar{r}^H(n) \bar{A} \bar{A}^H \bar{r}(n)} \quad (8.126)$$

Thus, the procedure to use CGM is to find the residual and the corresponding weights and update until convergence is satisfied. It can be shown that the true solution can be found in no more than K iterations. This condition is known as *quadratic convergence*.

Example 8.13 For the $M = 8$ -element array with elements a half-wavelength apart, find the optimum weights under the following conditions: The arriving signal of interest is $s = \cos(\pi k/K)$; arriving at an angle of 45° . One interfering signal is defined as $I_1 = \text{randn}(1,K)$; arriving at -30° . The other interfering signal is defined as $I_2 = \text{randn}(1,K)$; arriving at 0° . The noise has a variance of $\sigma^2 = .001$. Thus, $n = \sigma * \text{randn}(1,K)$. Use the CGM method to find the optimum weights when using a block size $K = 20$. Plot the norm of the residual for all iterations. Plot the resulting pattern.

Solution The MATLAB code `sa_ex8_13.m` is used to solve for the optimum weights using CGM. The plot of the norm of the residuals is given in Fig. 8.25. It can be seen that the residual drops to very small levels after 14 iterations. The plot of the resulting pattern is shown in Fig. 8.26. It can be seen that two nulls are placed at the two angles of arrival of the interference.

It should be especially noted that MATLAB provides a conjugate gradient function that can solve for the optimum weights. The function is the least squares implementation of CG and is given by $w = \text{lsqr}(A,d)$; A is the matrix as defined earlier and d is the desired signal vector containing K -data samples. Using the parameters in Example 8.12, we can use the MATLAB function `lsqr(A,d)` and produce the array factor plot in Fig. 8.27. The MATLAB function produces nulls in the correct locations but the sidelobe levels are higher than produced by the code `sa_ex8_12.m`.

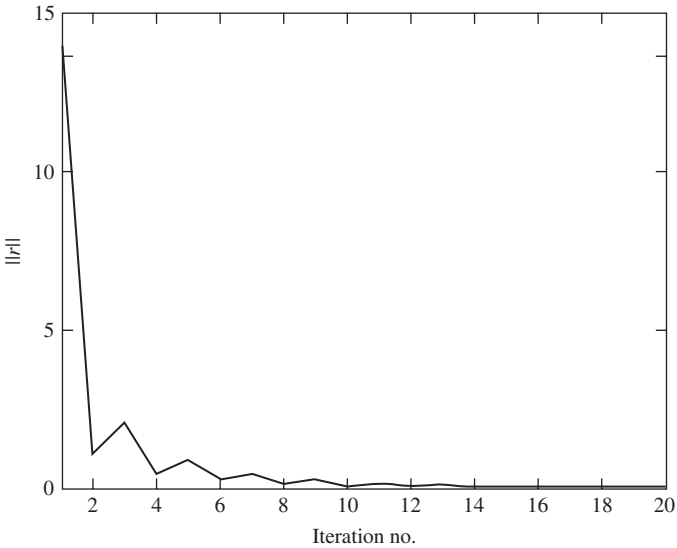


Figure 8.25 Norm of the residuals for each iteration.

8.4.7 Spreading Sequence Array Weights

A radically different approach to wireless beamforming has been proposed which can augment or replace traditional approaches to DBF or adaptive arrays. This new approach does not perform electronic or digital phase shifting. It does not rely on adaptive methods; however,

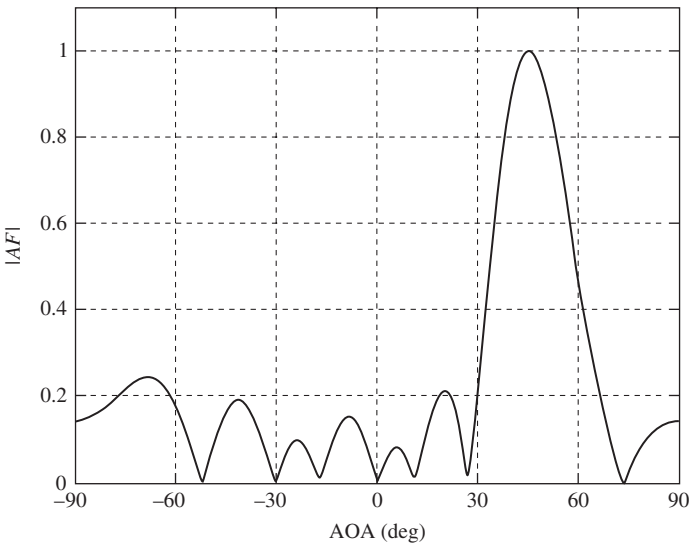


Figure 8.26 Array pattern using CGM.

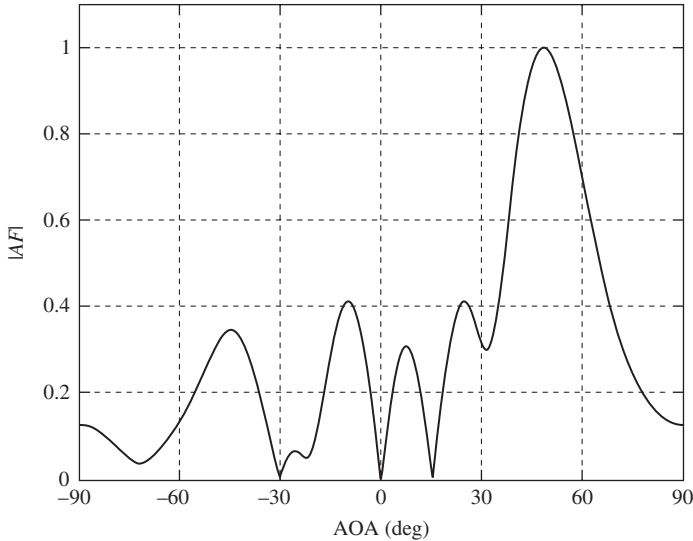


Figure 8.27 Array pattern resulting from use of the MATLAB CGM.

receive beams are produced in directions of interest. An array correlation matrix is not computed, array weights are not adapted, an a priori knowledge of the signal environment is not required, and a desired or pilot signal is not used. Rather, a clever spreading technique is used that uniquely defines all received directions of interest. This algorithm can create virtual contiguous beams like a Butler matrix, or it can selectively look in any specific directions of interest.

This new DBF receiver is based upon a radical shift in traditional DBF methods. This *out of the box* approach blends the virtues of switched beam and adaptive array technologies while avoiding many of their respective weaknesses. This solution does not direct beams by any previously known method, although this approach falls under the general topic of *waveform diversity*. This new method provides the same spatial resolution as traditional arrays while simultaneously looking in all desired directions of interest.

The previous adaptive methods discussed rely on subspace methods, steepest descent methods, gradients, blind adaptive algorithms, signal coherence, constant moduli, and other known signal properties in order to provide feedback to control the array weights. There are many drawbacks to the previous methods including computational intensity, a required knowledge of signal statistics, signal independence, and slow convergence rates.

This wholly new approach is based upon a recent patent by Carl Elam [53]. In this new method, adaptation is not necessary and

tracking is not performed, while excellent results are achieved. This approach can use any antenna configuration and, within the resolution of the antenna geometry, can simultaneously receive numerous signals of unequal strength at various angles. This novel approach can use any arbitrary two-dimensional array to produce an instantaneous pin-cushion array.

The new technique essentially works by designing the array weights to be time-varying random phase functions. The weights act to modulate each array element output. Specifically, the array outputs are weighted or modulated with a set of statistically independent polyphase chipping sequences. A different and independent modulating waveform is used for each antenna output. The chips are the same as chips defined in the traditional communications sense. The phase modulation waveforms purposely shred the phase of each antenna output by chipping at rates much higher than the baseband frequency of the message signal. This shredding process breaks up the phase relationship between all array elements and thus purposely eliminates array phase coherence. This is opposite of the traditional goal to achieve phase coherence for a specific look angle. The receiver then can *see* all incoming signals simultaneously without the necessity of steering or adapting because the array elements, for the moment, become statistically independent from one another. The chipped incoming waveforms are processed in a quadrature receiver and are subsequently compared to similar chipped waveforms stored in memory. The memory waveforms are created based upon expected angles of arrival. The theoretical spatial capacity is simply the angular space of interest divided by the antenna array beam solid angle.

8.4.8 Description of the new SDMA receiver

The novel DBF receiver is depicted in Fig. 8.28. The novelty of this new approach is vested in the nature of the signals $\beta_n(t)$, the unique array signal memory, and correlation based detection.

The new SDMA digital beamformer can be used with any arbitrary N -element antenna array. It can be a linear array but should preferably be a two-dimensional or three-dimensional random array such that the antenna geometry and element phasing is unique for each incoming angle-of-arrival. Ideally the receive array should have an equal angular resolution over all receive angles of interest. For purposes of illustration, the array used in this discussion will be an N -element linear array.

The incoming signals arrive at angles θ_ℓ where $\ell = 1, 2 \dots L$. Each different angle of arrival produces a unique array element output with a unique phase relationship between each element. These phase relationships will be used in conjunction with the modulations $\beta_n(t)$ to produce a unique summed signal $y^r(t)$.

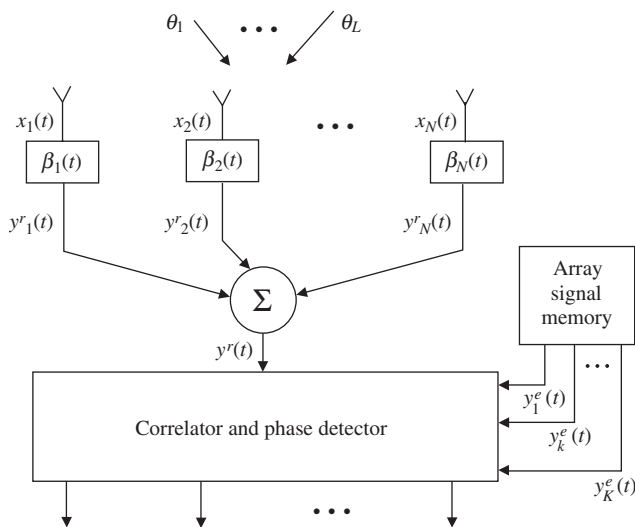


Figure 8.28 SDMA quadrature receiver.

Corresponding to the actual N -element antenna array is a second virtual array modeled in memory. The virtual array is modeled after the actual physical array used. The memory array has K possible virtual outputs for each expected direction $\theta_k(k = 1, 2, \dots, K)$. The total number of expected directions K should be less than or equal to the number of antenna elements N . These memory signals are generated based upon a knowledge of the antenna array geometry and the calculated phase delays for each specific direction of interest. The expected directions are generally chosen by the user to be no closer than the angular resolution allowed by the array. Consideration can be given to the local topography and potential multipath directions in eliminating possible unwanted expected directions. All potential incoming directions may not be used if it is desired to block interfering signals from known directions.

The N -element array outputs and the N -antenna array memory outputs are both *phase modulated* (PM) by the same set of N pseudo-noise (*pn*) phase modulating sequences. The n th phase modulating sequence will be designated as $\beta_n(t)$. $\beta_n(t)$ is composed of M polyphase chips. Each chip is of length τ_c and the entire sequence is of length $T = M\tau_c$. The chip rate is chosen to be much greater than the Nyquist rate of the incoming baseband signal modulation. The purpose for this *over sampling* is so that the phase modulation of the incoming signal is nearly constant over the entire group of M chips. In general, the goal should be such that $T \leq 1/(4B_m)$ where B_m is the message signal bandwidth.

Each phase modulating waveform $\beta_n(t)$ is used to modulate or *tag* each array output with a unique marking or identifying waveform. $\beta_n(t)$

deliberately shreds or scrambles the phase of the signal at array element n . This shredding process, temporarily scrambles the phase relationship between all other array elements. The desired element phasing is restored in the correlator if the incoming signal correlates with one of the memory signals.

Examples of some usable pn modulating sequences are Gold, Kasami, Welty, Golay, or any other sequences with similar statistical and orthogonal properties [54], [56]. These properties will aid in the identification of the exact directions of arrival of the L incoming signals. Figure 8.29 shows a typical example of the first two of the set of N bi-phase sequences applied to an N -element array. The total code length $TB_m = .25$.

Since the phase out of each antenna element is intentionally scrambled by the modulating sequences, the array pattern is randomized during each chip in the sequence. The momentarily scrambled pattern changes every τ_c seconds. As an example, the array patterns are plotted for the first four chips of an $N = 10$ -element array as shown in Fig. 8.30. Since the phase relationship between all elements is scrambled, the array patterns for each new set of chips is random. In the limit, as the number of chips increases, the average array pattern, over all chips, becomes uniform.

Each baseband output of the receive array will have a complex voltage waveform whose phase will consist of each emitter's message signal $m_\ell(t)$ and the unique receive antenna element phase contributions.

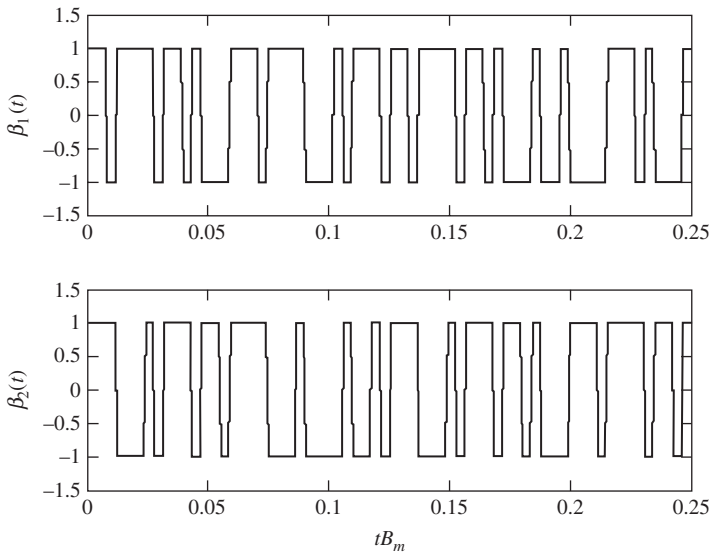


Figure 8.29 Two sample pn bi-phase sequences.

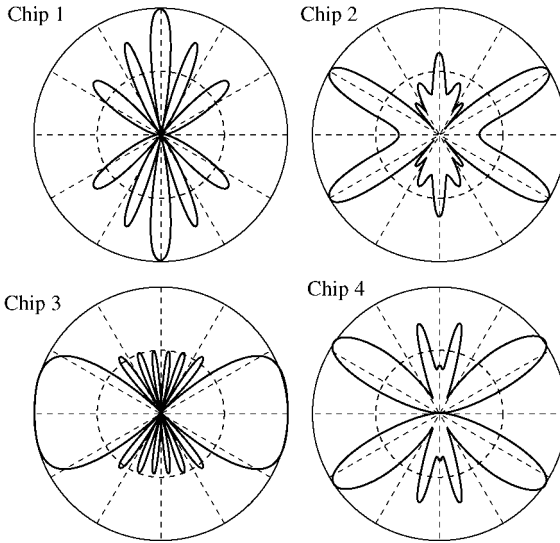


Figure 8.30 Scrambled patterns for the first four chips.

Ignoring the space loss and polarization mismatches, the received baseband array output is given in vector form as

$$\begin{aligned} \bar{x}^r(t) &= \begin{bmatrix} 1 & \cdots & 1 \\ e^{jkd \sin(\theta_1)} & \cdots & e^{jkd \sin(\theta_L)} \\ \vdots & \cdots & \vdots \\ e^{j(n-1)kd \sin(\theta_1)} & \cdots & e^{j(n-1)kd \sin(\theta_L)} \end{bmatrix} \begin{bmatrix} e^{jm_1(t)} \\ \vdots \\ e^{jm_L(t)} \end{bmatrix} \\ &= \bar{A}^r \cdot \bar{s}^r(t) \end{aligned} \quad (8.127)$$

where $m_\ell(t)$ = ℓ th emitter's phase modulation

d = array element spacing

k = wavenumber

θ_ℓ = angle of arrival of the ℓ th incoming signal

$\bar{a}_\ell^r = [1, e^{jkd \sin(\theta_\ell)}, \dots, e^{j(n-1)kd \sin(\theta_\ell)}]^T$

= steering vector for direction θ_ℓ

$\bar{A}^r$ = matrix of steering vectors for all angles of arrival θ_ℓ

$\bar{s}^r$ = vector of arriving signal baseband phasors

The received signals, for each array output, are phase modulated with the chipping sequences as described earlier. The chipping waveforms can be viewed as phase-only array weights. These array weights can be depicted as the vector $\bar{\beta}(t)$. The total weighted or chipped array output is called the *received signal vector* and is given by

$$y^r(t) = \bar{\beta}(t)^T \cdot \bar{x}^r(t) \quad (8.128)$$

In a similar way, the array signal memory steering vectors are created based upon M expected angles-of-arrival θ_m .

$$\begin{aligned}\bar{A}^e &= \begin{bmatrix} 1 & \cdots & 1 \\ e^{jkd \sin(\theta_1)} & \cdots & e^{jkd \sin(\theta_K)} \\ \vdots & \cdots & \vdots \\ e^{j(n-1)kd \sin(\theta_1)} & \cdots & e^{j(n-1)kd \sin(\theta_K)} \end{bmatrix} \\ &= [\bar{a}_1^e \quad \cdots \quad \bar{a}_K^e]\end{aligned}\quad (8.129)$$

where $\bar{a}_k^e$ is the steering vector for expected direction θ_k and $\bar{A}^e$ is the matrix of steering vectors for expected direction θ_k .

The memory has K outputs, one for each expected direction θ_k . Each memory output, for the expected angle θ_k is given by

$$y_k^e(t) = \bar{\beta}(t)^T \cdot \bar{a}_k^e \quad (8.130)$$

The signal correlator is designed to correlate the actual received signal with the conjugate of the various expected direction memory signals. This is similar to matched filter detection. The best correlations occur when the actual AOA matches the expected AOA. The correlation can be used as a discriminant for detection. Since the arriving signals have a random arrival phase delay, a quadrature correlation receiver should be employed such that the random carrier phase does not affect the detection (Haykin [54]). The general complex correlation output, for the k th expected direction, is given as

$$R_k = \int_t^{t+T} y^r(t) \cdot y_k^{e*}(t) dt = |R_k| e^{j\phi_k} \quad (8.131)$$

where R_k is the correlation magnitude at expected angle θ_k and ϕ_k is the correlation phase at expected angle θ_k .

The new SDMA receiver does not process the incoming signals with phase shifters or beam steering. We do not *look* for the emitter direction by steering but actually find the direction by correlation. The correlation magnitude $|R_k|$ is used as the discriminant to determine if a signal is present at the expected angle θ_k . If the discriminant exceeds a predetermined threshold, a signal is deemed present and the phase is calculated. Since it is assumed that the emitter PM is nearly constant over the code length $M\tau_c$, the correlator output phase angle is approximately the average of the emitter PM. Thus

$$\phi_k = \arg(R_k) \approx \tilde{m}_k \quad (8.132)$$

where

$$\tilde{m}_k = \frac{1}{T} \int_t^{t+T} m_k(t) dt = \text{average of emitter's modulation at angle } \theta_k$$

The average phase $\tilde{m}_k$ retrieved for each quarter cycle of the $m(t)$ can be used to reconstruct the user's phase modulation using an FIR filter.

Example using bi-phase chipping. We can let all chipping waveforms, $\beta_n(t)$, be defined as *bi-phase pn sequences*. It is instructive to show the appearance of the signals at various stages in the receiver. Figure 8.31 shows the n th element received baseband modulation, the n th phase modulating waveform $\beta_n(t)$, the n th phase modulated output $y_n^r(t)$, and finally the N combined phase modulated outputs $y^r(t)$.

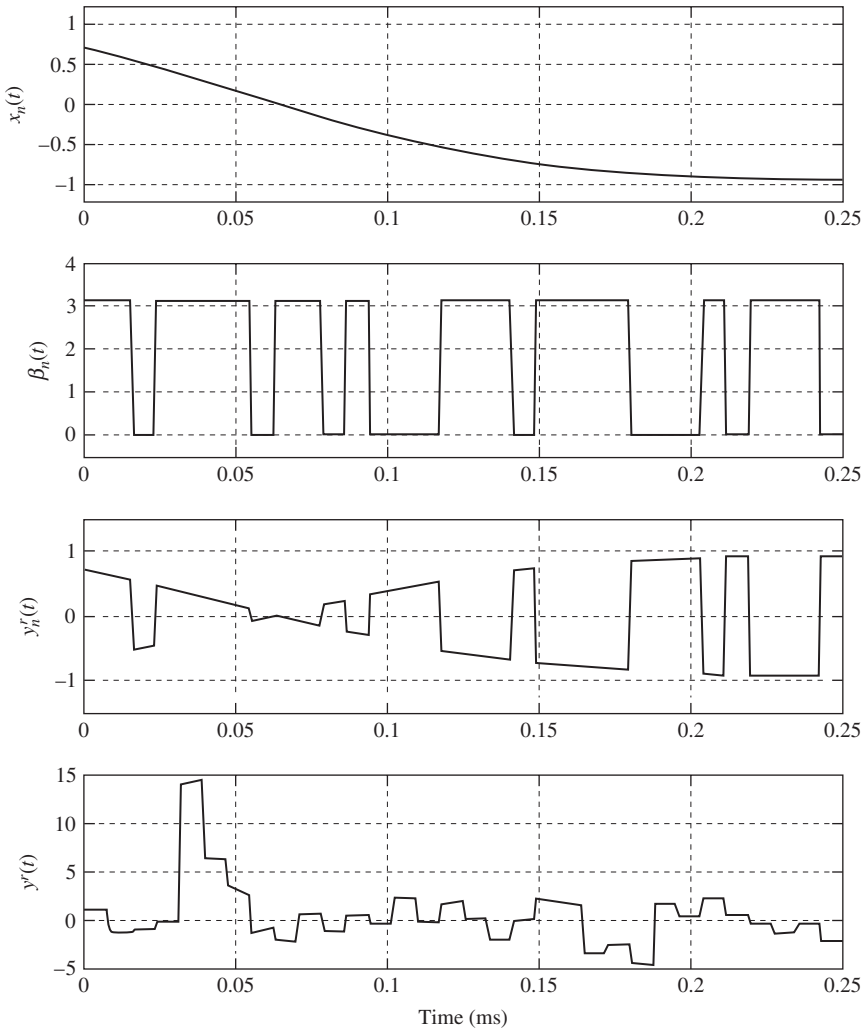


Figure 8.31 Display of waveforms throughout the receiver.

The fourth waveform is obviously unique owing to the arrival phase angle of the incoming signal from direction θ_i and the chipping by the N independent phase modulating waveforms. This specific chipped and summed waveform is unique to the direction of arrival and can be correlated with pre-calculated waveforms stored in memory.

Figure 8.32 compares the correlation magnitude in Eq. (8.131) for an incoming signal ($\theta = 0^\circ$) superimposed with the array factor for a typical N -element linear array. The following receiver values are used: $N = 21, d = \lambda/2, M = 128$ chips. Figure 8.32 demonstrates that the new SDMA receiver achieves the same angular resolution as the conventional linear array provided that the number of chips M is adequately large. Shorter sequences can be used when the sequences are not only independent but also orthogonal.

When multiple signals arrive at the receive antenna array, they can be detected through the same process of correlating all incoming signals with the expected signals stored in memory. Figure 8.33 demonstrates the correlation for five arriving signals at equally spaced angles of $-60^\circ, -30^\circ, 0^\circ, 30^\circ, 60^\circ$. The chipping sequences are binary Welty codes of length 32.

After a signal is deemed present near the expected angle θ_k , we can now find the average phase modulation ϕ_k defined in Eq. (8.133). Figure 8.34 shows an example where sequences of M chips are used four times in succession to reconstruct a piecewise estimate of the message

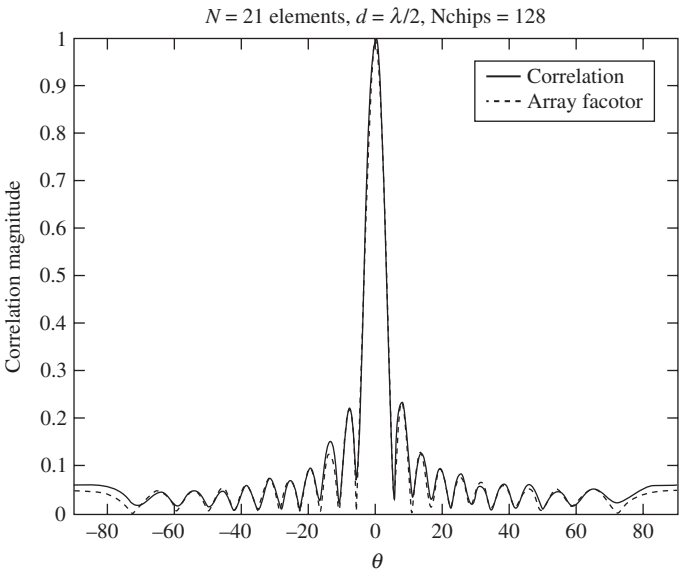


Figure 8.32 Comparing the new SDMA receiver to a traditional linear array.

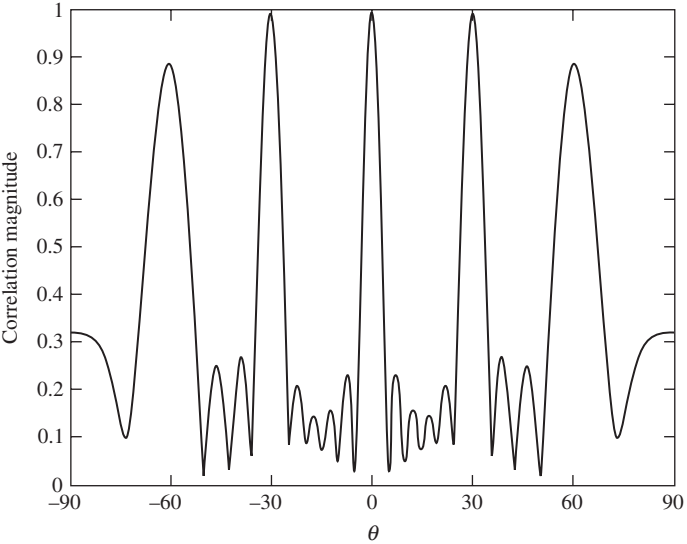


Figure 8.33 Correlation for five equally spaced arrival angles.

signal $m_k(t)$. Figure 8.34 shows the original PM $m_k(t)$ and the estimate calculated for each quarter cycle.

It can be seen that the predicted average phase for each quarter cycle can be fed into an FIR filter to reconstruct the original emitter modulation.

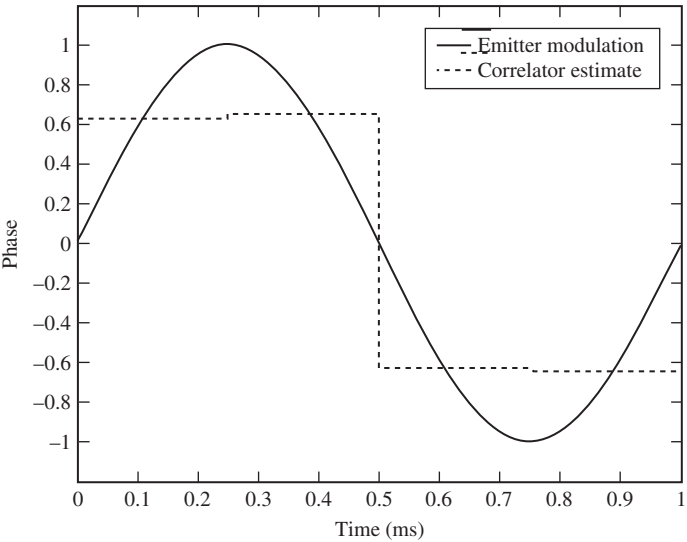


Figure 8.34 Emitter modulation with receiver estimate.

In summary, the new SDMA receiver uses spreading sequences, as array weights, to shred the phase out of each array element. This approach provides a radical and novel alternative to traditional fixed beam and adaptive array approaches. The spreading sequences are composed of polyphase codes that have chip rates much greater than the anticipated emitter baseband modulation. The shredding momentarily randomizes the phase relationship between each element output. This in turn causes the pattern to be randomized during each chip of the spreading sequence. Over the M chip duration of the spreading sequence, the average pattern for the array approaches that of an omnidirectional array. The shredded signal is reconstructed with a quadrature correlation receiver by correlating with memory signals. The memory signals are created by a similar shredding process but for expected angles of arrival. Correlator outputs are generated for each expected AOA. The correlation magnitude is used as a discriminant to decide if a signal is present at the expected angle θ_k . The correlation phase angle is a piecewise estimate of the original emitter's modulation.

This new approach has numerous advantages over traditional SDMA array antennas. The receiver is superior to a switched beam array because contiguous beams can be formed in any regions of interest without the need for hardware phase shifters. The beams are created by correlation. The receiver memory can simply be altered to redirect the beams to new regions of interest if so desired. The elimination of phase shifters can result in a significant cost savings.

The new receiver also has advantages over adaptive arrays because it does not require adaptation, it can process multiple angles of arrival simultaneously, and it is not limited by acquisition or tracking speeds. Interfering signals are minimized because the shredded waveforms at interferer angles do not correlate well with expected direction waveforms stored in memory.

Any arbitrary and/or random antenna array geometry can be incorporated into this new approach as long as the expected memory signals are based upon the array geometry. The array resolution achieved is consistent with the limitations of the array geometry.

References

1. Barton, P., "Digital Beamforming for Radar," *IEEE Proceedings on Pt. F*, Vol. 127, pp. 266–277, Aug. 1980.
2. Brookner, E., "Trends in Array Radars for the 1980s and Beyond," *IEEE Antenna and Propagation Society Newsletter*, April 1984.
3. Steyskal, H., "Digital Beamforming Antennas—An Introduction," *Microwave Journal*, pp. 107–124, January 1987.

4. Skolnik, M., "System Aspects of Digital Beam Forming Ubiquitous Radar," Naval Research Lab, Report No. NRL/MR/5007—02-8625, June 28, 2002.
5. Skolnik, M., *Introduction to Radar Systems*, 3d ed. McGraw-Hill, New York, 2001.
6. Curtis, T., "Digital Beamforming for Sonar Systems," *IEEE Proceedings on Pt. F*, Vol. 127, pp. 257–265, Aug. 1980.
7. Litva, J., and T. Kowk-Yeung Lo, *Digital Beamforming in Wireless Communications*, Artech House, 1996.
8. Liberti, J., and T. Rappaport, *Smart Antennas for Wireless Communications: IS-95 and Third Generation CDMA Applications*, Prentice Hall, New York, 1999.
9. Van Atta, L., "Electromagnetic Reflection," U.S. Patent 2908002, Oct. 6, 1959.
10. Margerum, D., "Self-Phased Arrays," in *Microwave Scanning Antennas*, Vol. 3, *Array Systems*, Ch. 5. ed. Hansen, R.C. Academic Press, New York, 1966.
11. York, R., and T. Itoh, "Injection and Phase-Locking Techniques for Beam Control," *IEEE Transactions on MTT*, Vol. 46, No. 11, pp. 1920–1920, Nov. 1998.
12. Howells, P., "Intermediate Frequency Sidelobe Canceller," U. S. Patent 3202990, Aug. 24, 1965.
13. Howells, P., "Explorations in Fixed and Adaptive Resolution at GE and SURC," *IEEE Transactions on Antenna and Propagation, Special Issue on Adaptive Antennas*, Vol. AP-24, No. 5, pp. 575–584, Sept. 1976.
14. Applebaum, S., "Adaptive Arrays," Syracuse University Research Corporation, Rep. SPL TR66-1, August 1966.
15. Applebaum, S., "Adaptive Arrays," *IEEE Transactions on Antenna and Propagation*, Vol. AP-24, No. 5, pp. 585–598, Sept. 1976.
16. Widrow, B., and M. Hoff, "Adaptive Switch Circuits," IRE Wescom, Convention Record, Part 4, pp. 96–104, 1960.
17. Widrow, B., P. Mantey, L. Griffiths, et al., "Adaptive Antenna Systems," *Proceedings of the IEEE*, Vol. 55, Dec. 1967.
18. Gabriel, W., "Adaptive Processing Antenna Systems," *IEEE Antenna and Propagation Newsletter*, pp. 5–11, Oct. 1983.
19. Godara, L., "Application of Antenna Arrays to Mobile Communications, Part II: Beam-Forming and Direction-of-Arrival Considerations," *Proceedings of the IEEE*, Vol. 85, No. 8, pp. 1195–1245, Aug. 1997.
20. Reed, I., J. Mallett, and L. Brennen, "Rapid Convergence Rate in Adaptive Arrays," *IEEE Transactions on Aerospace and Electronics Systems*, Vol. AES-10, pp. 853–863, Nov. 1974.
21. Capon, J., "High-Resolution Frequency-Wavenumber Spectrum Analysis," *Proceedings of the IEEE*, Vol. 57, No. 8, pp. 1408–1418, Aug. 1969.
22. Makhoul, J., "Linear Prediction: A Tutorial Review," *Proceedings of the IEEE*, Vol. 63, pp. 561–580, 1975.
23. Burg, J.P., "The Relationship between Maximum Entropy Spectra and Maximum Likelihood Spectra," *Geophysics*, Vol. 37, pp. 375–376, April 1972.
24. Burg, J.P., "Maximum Entropy Spectrum Analysis," Ph.D. Dissertation, Dept. of Geophysics, Stanford University, Stanford CA, 1975.
25. Pisarenko, V.F., "The Retrieval of Harmonics from a Covariance Function," *Geophysical Journal of the Royal Astronomical Society*, Vol. 33, pp. 347–366, 1973.
26. Reddi, S.S., "Multiple Source Location—A Digital Approach," *IEEE Transactions on AES*, Vol. 15, No. 1, Jan. 1979.
27. Kumaresan, R., and D. Tufts, "Estimating the Angles of Arrival of Multiple Plane Waves," *IEEE Transactions on AES*, Vol. AES-19, pp. 134–139, 1983.
28. Schmidt, R., "Multiple Emitter Location and Signal Parameter Estimation," *IEEE Transactions on Antenna Propagation*, Vol. AP-34, No. 2, pp. 276–280, March 1986.
29. Roy, R., and T. Kailath, "ESPRIT—Estimation of Signal Parameters Via Rotational Invariance Techniques," *IEEE Transactions on ASSP*, Vol. 37, No. 7, pp. 984–995, July 1989.
30. Monzingo, R., and T. Miller, *Introduction to Adaptive Arrays*, Wiley Interscience, John Wiley & Sons, New York, 1980.

31. Harrington, R., *Field Computation by Moment Methods*, MacMillan, New York, Chap. 10, pp. 191, 1968.
32. Van Trees, H., *Detection, Estimation, and Modulation Theory: Part I*, Wiley, New York, 1968.
33. Van Trees, H., *Optimum Array Processing, Part IV of Detection, Estimation, and Modulation Theory*, Wiley Interscience, New York, 2002.
34. Haykin, S., H. Justice, N. Owsley, et al., *Array Signal Processing*, Prentice Hall New York, 1985.
35. Cohen, H., *Mathematics for Scientists and Engineers*, Prentice Hall, New York, 1992.
36. Godara, L., *Smart Antennas*, CRC Press, Boca Raton, FL, 2004.
37. Haykin, S., *Adaptive Filter Theory*, 4th ed., Prentice Hall, New York, 2002.
38. Golub, G.H., and C.H. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, 3d ed., 1996.
39. Godard, D.N., "Self-Recovering Equalization and Carrier Tracking in Two-Dimensional Data Communication Systems," *IEEE Transactions on Communications*, Vol. Com-28, No. 11, pp. 1867–1875, Nov. 1980.
40. Larimore, M., and J. Treichler, "Convergence Behavior of the Constant Modulus Algorithm, Acoustics," *IEEE International Conference on ICASSP '83*, Vol. 8, pp. 13–16, April 1983.
41. Treichler, J., and B. Agee, "A New Approach to Multipath Correction of Constant Modulus Signals," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. Assp-31, No. 2, pp. 459–472, April 1983.
42. Agee, B., "The Least-Squares CMA: A New Technique for Rapid Correction of Constant Modulus Signals," *IEEE International Conference on ICASSP '86*, Vol. 11, pp. 953–956, April 1986.
43. Sorenson, H., "Least-Squares Estimation: From Gauss to Kalman," *IEEE Spectrum*, Vol. 7, pp. 63–68, July 1970.
44. Rong, Z., "Simulation of Adaptive Array Algorithms for CDMA Systems," Master's Thesis MPRG-TR-96-31, Mobile & Portable Radio Research Group, Virginia Tech, Blacksburg, VA, Sept. 1996.
45. Stoica, P., and R. Moses, *Introduction to Spectral Analysis*, Prentice Hall, New York, 1997.
46. Hestenes, M., and E. Stiefel, "Method of Conjugate Gradients for Solving Linear Systems," *Journal of Research of the National Bureau of Standards*, Vol. 49, pp. 409–436, 1952.
47. Fletcher, R., and M. Powell, "A Rapidly Convergent Descent Method for Minimization," *Computer Journal*, Vol. 6, pp. 163–168, 1963.
48. Fletcher, R., and C. Reeves, "Function Minimization by Conjugate Gradients," *Computer Journal*, Vol. 7, pp. 149–154, 1964.
49. Choi, S., *Application of the Conjugate Gradient Method for Optimum Array Processing*, Book Series on PIER (Progress in Electromagnetics Research), Vol. 5, Elsevier, Amsterdam, 1991.
50. Choi, S., and T. Sarkar, Adaptive Antenna Array Utilizing the Conjugate Gradient Method for Multipath Mobile Communication, *Signal Processing*, Vol. 29, pp. 319–333, 1992.
51. Godara, L., *Smart Antennas*, CRC Press, Boca Raton, FL, 2004.
52. Sadiku, M., *Numerical Techniques in Electromagnetics*, 2d ed., CRC Press, Boca Raton, FL, 2001.
53. Elam, C., "Method and Apparatus for Space Division Multiple Access Receiver," Patent No. 6,823,021, Rights assigned to Greenwich Technology Associates, One Soundview way, Darien, CT.
54. Simon, H. *Communication Systems*, 2d ed., p. 580. Wiley, New York, 1983.
55. Ziemer, R.E., R. L. Peterson, *Introduction to Digital Communication*, Prentice Hall, pp. 731–742, 2001.
56. Skolnik, M. *Radar Handbook*, 2nd ed., McGraw-Hill, pp. 10.17–10.26, 1990.

Problems

8.1 For an $M = 5$ -element array with $d = \lambda/2$, the desired signal arrives at $\theta = 20^\circ$, one interferer arrives at -20° , while the other interferer arrives at $+45^\circ$. The noise variance is $\sigma_n^2 = .001$. Use the Godara method outlined in Eq. (8.8)

- What are the array weights?
- Plot the magnitude of the weighted array pattern for $-90^\circ < \theta < 90^\circ$.

8.2 Maximum SIR method: Given a $M = 5$ -element array with spacing $d = .5\lambda$ and noise variance $\sigma_n^2 = .001$, a desired received signal arriving at $\theta_0 = 20^\circ$, and two interferers arrive at angles $\theta_1 = -30^\circ$ and $\theta_2 = -45^\circ$, assume that the signal and interferer amplitudes are constant and that $\bar{R}_{ss} = \bar{R}_{ii} = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix}$.

- Use MATLAB to calculate $\text{SIR}_{\max}$
- What are the normalized weights?
- Plot the resulting pattern.

8.3 Minimum Mean Square Error (MSE) method: For an $M = 2$ -element array with elements a half-wavelength apart, find the optimum Wiener weights under the following conditions: The arriving signal of interest is $s(t) = e^{j\omega t}$ arriving at the arbitrary angle of θ_s . The noise is zero-mean Gaussian with arbitrary variance σ_n^2 . Allow the desired signal $d(t) = s(t)$.

- What is the signal correlation matrix, $\bar{R}_{ss}$, in symbolic form?
- What is the array correlation matrix, $\bar{R}_{xx}$, in symbolic form?
- What is the correlation vector $\bar{r}$?
- What is the correlation matrix inverse, $\bar{R}_{xx}^{-1}$, in symbolic form?
- Symbolically derive the equation for the weights.
- What are the exact weights for $\theta_s = 30^\circ$ and for $\sigma_n^2 = .1$?

8.4 Minimum Mean Square Error (MMSE) method: Let the array be a two element array where the arriving signal is coming in at $\theta = 0^\circ$. Let the desired signal equal the arriving signal with $s = 1$. Thus, $\bar{x} = \begin{bmatrix} 1 \\ e^{jhd \sin(\theta)} \end{bmatrix}$. Let $d = \lambda/2$.

- Derive the performance surface using Eq. (8.23) and plot for $-4 < w_1, w_2 < 4$.
- Derive the solutions for w_1 and w_2 by using $\frac{\partial}{\partial w_1}$ and $\frac{\partial}{\partial w_2}$.
- How is the derived solution consistent with the plot? (Recall that the $\nabla_w E[|\varepsilon|^2] = 0$ gives the *bottom* of the curve.)
- Derive the performance surface using Eq. (8.23) for the arrival angle of $\theta = 30^\circ$
- Derive the solutions for w_1 and w_2 by using $\frac{\partial}{\partial w_1}$ and $\frac{\partial}{\partial w_2}$ for the angle given in part d).

8.5 MMSE method: Given an $M = 5$ -element array with spacing $d = .5\lambda$, a received signal energy $S = 1$ arriving at $\theta_0 = 30^\circ$, and two interferers arriving at angles $\theta_1 = -20^\circ$ and $\theta_2 = 40^\circ$, with noise variance $\sigma_n^2 = .001$.

Assume that the signal and interferer amplitudes are constant and that $\bar{R}_{ss} = \bar{R}_{ii} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

- (a) Use MATLAB to calculate the optimum weights.
- (b) Plot the resulting pattern.

8.6 Maximum likelihood (ML) method: Given an $M = 5$ -element array with spacing $d = .5\lambda$ which has a received signal arriving at the angle $\theta_0 = 45^\circ$, with noise variance $\sigma_n^2 = .01$.

- (a) Use MATLAB to calculate the optimum weights
- (b) Plot the magnitude of the weighted array pattern for $-90^\circ < \theta < 90^\circ$.

8.7 Minimum variance (MV) method: Given an $M = 5$ -element array with spacing $d = .5\lambda$ which has a received signal arriving at the angle $\theta_0 = 40^\circ$, one interferer arriving at -20° , and noise with a variance $\sigma_n^2 = .001$.

- (a) Use MATLAB to calculate the optimum weights
- (b) Plot the magnitude of the weighted array pattern for $-90^\circ < \theta < 90^\circ$.

8.8 Least mean squares (LMS) method: For an $M = 2$ -element array with elements a half-wavelength apart, find the LMS weights using Eqs. (8.48) and (8.56) under the following conditions: $\mu = .5$, $\sigma_n^2 = 0$, the arriving signal is $s(t) = 1$, the angle of arrival is $\theta_s = 45^\circ$, the desired signal $d(t) = s(t)$. Set the initial array weights such that $\bar{w}(1) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. There are no other arriving signals.

- (a) By hand calculate the array weights for the next three iterations (i.e., $\bar{w}(2), \bar{w}(3), \bar{w}(4)$)
- (b) What is the error $|\varepsilon(k)|$ for $k = 2, 3, 4$?
- (c) Use MATLAB to calculate the weights and error for 20 iterations. Plot the absolute value of each weight vs. iteration k on one plot and the absolute value of the error vs. iteration k on another plot.

8.9 LMS method: Given an $M = 8$ -element array with spacing $d = .5\lambda$ which has a received signal arriving at the angle $\theta_0 = 40^\circ$, an interferer at $\theta_1 = -20^\circ$. Assume that the desired received signal vector is defined by $\bar{x}_s(k) = \bar{a}_0 s(k)$ where $s(k) = \sin(\pi t(k)/T)$; with $T = 1$ ms and $t = (1 : 100) * T / 100$. Assume the interfering signal vector is defined by $\bar{x}_i(k) = \bar{a}_1 i(k)$ where $i(k) = \text{randn}(1, 100)$. Both signals are nearly orthogonal over the time interval T . Let the desired signal $d(k) = s(k)$. Assume that the initial array weights are all zero. Allow for 100 iterations. Let step size $\mu = .02$.

- (a) Calculate the eight array weights for 100 iterations.
- (b) Plot the magnitude of the each weight vs. iteration number on the same plot.
- (c) Plot the mean square error $|e|^2$ vs. iteration number
- (d) What are the array weights at the 100th iteration?
- (e) Plot the magnitude of the weighted array pattern for $-90^\circ < \theta < 90^\circ$.

8.10 Sample matrix inversion (SMI) method: Use the SMI method to find the weights for an $M = 8$ -element array with spacing $d = .5\lambda$. Let the received signal arrive at the angle $\theta_0 = 45^\circ$. One interferer arrives at $\theta_1 = -45^\circ$. Assume that the desired received signal vector is defined by $\bar{x}_s(k) = \bar{a}_0 s(k)$ where $s(k) =$

$\sin(2 * \pi * t(k)/T)$ with $T = 2$ ms. Let the block length be $K = 50$. Time is defined as $t = (1 : K) * T/K$. Assume the interfering signal vector is defined by $\bar{x}_i(k) = \bar{a}_1 i(k)$ where $i(k) = \text{randn}(1, K) + \text{randn}(1, K)j$. Let the desired signal $d(k) = s(k)$. In order to keep the correlation matrix inverse from becoming singular, add noise to the system with variance $\sigma_n^2 = .01$. Define the noise in MATLAB as $n = \text{randn}(N, K) * \text{sqrt}(\text{sig2})$;

(a) Find the weights.

(b) Plot the magnitude of the weighted array pattern for $-90^\circ < \theta < 90^\circ$.

8.11 Repeat Prob. 8 for $M = 5$, $d = .5\lambda$, $\theta_0 = 30^\circ$, $\theta_1 = -20^\circ$, and allow the received signal to be a phase modulation such that $s(k) = \exp(1j * .5 * \pi * \sin(\pi * t(k)/T))$;

8.12 Recursive least squares (RLS) method: Let the array be an $M = 7$ -element array with spacing $d = .5\lambda$ with a received signal arriving at the angle $\theta_0 = 30^\circ$, an interferer at $\theta_1 = -20^\circ$, and the other interferer at $\theta_2 = -40^\circ$. Use MATLAB to write an RLS routine to solve for the desired weights. Assume that the desired received signal vector is defined by $\bar{x}_s(k) = \bar{a}_0 s(k)$ where $s(k) = \exp(1j * \sin(\pi * t(k)/T))$; with $T = 1$ ms. Let there be $K = 50$ time samples such that $t = (0 : K - 1) * T/(K - 1)$. Assume that the interfering signals $i1(k) = \exp(1j * \pi * \text{randn}(1, K))$; $i2(k) = \exp(1j * \pi * \text{randn}(1, K))$; (MATLAB changes the random numbers each time so that $i1$ and $i2$ are different. Let the desired signal $d(k) = s(k)$. In order to keep the correlation matrix inverse from becoming singular, add noise to the system with variance $\sigma_n^2 = .01$. Initialize all array weights as zero. Set the forgetting factor $\alpha = .995$.

(a) Plot the magnitude of the first weight for all iterations.

(b) Plot the magnitude of the weighted array pattern for $-90^\circ < \theta < 90^\circ$.

8.13 Constant modulus algorithm (CMA) method: Allow the same constant modulus signal to arrive at the receiver via a direct path and two additional multipaths and assume that the channel is frequency selective. Let the direct path arriving signal be defined as a 32-chip binary sequence where the chip values are ± 1 and are sampled four times per chip (see Example 8.10). The direct path signal arrives at 30° . The first multipath signal arrives at 0° but is 30 percent of the direct path in amplitude. The second multipath signal arrives at 20° but is 10 percent of the direct path in amplitude. Because of multipath, there will be slight time delays in the binary sequences causing dispersion. Let the dispersion be implemented by zero padding the signals. Zero pad the direct path by adding eight zeros to the back of the signal. Zero pad the first multipath signal with two zeros in front and two zeros in back. Zero pad the second multipath signal with four zeros up front. We will use the $P = 1$ CMA algorithm to define the optimum weights. Choose $\mu = .6$, $N = 6$ elements, and $d = \lambda/2$. Define the initial weights, $\bar{w}(1)$, to be zero. Plot the resulting pattern using MATLAB.

8.14 Least squares constant modulus algorithm (LS-CMA) method: Use the static version of the LS-CMA. Allow the same constant modulus signal to arrive at the receiver via a direct path and one additional multipath and assume that

the channel is frequency selective. Let the direct path arriving signal be defined as a 32-chip binary sequence where the chip values are ± 1 and are sampled four times per chip. Let the block length $K = 132$. The direct path signal arrives at 30° . A multipath signal arrives at -30° but is 50 percent of the direct path in amplitude. Because of multipath, there will be slight time delays in the binary sequences causing dispersion. Let the dispersion be implemented by zero padding the signals. Zero pad the direct path by adding four zeros to the back of the signal. Zero pad the multipath signal with two zeros in front and two zeros in back. Use the LS-CMA algorithm to find the optimum weights. $N = 9$ elements and $d = \lambda/2$. Add noise for each element in the array and let the noise variance be $\sigma_n^2 = .01$. Define the initial weights, $\bar{w}(1)$, to be one. Iterate three times. Plot the resulting pattern using MATLAB.

8.15 Repeat Prob. 14 but use the dynamic LS-CMA algorithm. Let the block length be $K = 22$. Allow the block to update for every iteration n . Stop after 5 iterations using 6 ($6 \cdot K = 132$) blocks. Plot the resulting pattern using MATLAB.

8.16 Conjugate gradient method (CGM): For the $M = 9$ -element array with elements a half-wavelength apart, find the optimum weights under the following conditions: The arriving signal of interest is $s = \sin(\pi k/K)$; arriving at an angle of 30° . One interfering signal is defined as $I_1 = \text{randn}(1, K)$; arriving at -20° . The other interfering signal is defined as $I_2 = \text{randn}(1, K)$; arriving at -45° . The noise has a variance of $\sigma_n^2 = .001$. Thus, $n = \sigma^* \text{randn}(1, K)$. Use the CGM to find the optimum weights when using a block size $K = 20$.

- (a) Plot the norm of the residual for all iterations.
- (b) Plot the resulting pattern using MATLAB.

Index

- adaptive
 - algorithms, 207–208
 - arrays, 207–209
 - beamforming, 227–251
- Ampere, A. M., 9
- analog-to-digital conversion, 2
- angle-of-arrival (AOA) estimation, 100, 169–201
 - Bartlett, 178
 - Capon, 179
 - ESPRIT, 197
 - linear prediction, 181
 - maximum entropy, 182
 - min-norm, 185
 - MUSIC, 187
 - Pisarenko harmonic decomposition, 183
 - root-min-norm, 195
 - root-MUSIC, 191
- angular spread, 146
- anisotropic, 43
- antenna field regions, 37–39
- antenna pattern, 43
- AOA estimation (*see* angle-of-arrival estimation)
- array
 - beamsteered, 74, 207
 - beamwidth, 47, 48
 - broadside, 73
 - end-fire, 73
 - even, 80
 - factor, 67
 - manifold vector, 69
 - odd, 81
 - processing, 210
 - propagation vector, 69
 - response vector, 69
 - steering vector, 69
 - vector, 69
- arrays, 66–101
 - beamsteered, 74, 207
 - circular, 89–91
 - linear, 65–79
 - planar, 91–93
- attenuation constant, 11
- autocorrelation, 117, 118
- autocorrelation-ergodic, 118
- Bartlett estimate, 178, 179
- Bartlett weights, 87
- beam solid angle, 49
- beamsteered arrays, 207
- beamsteered, 74
- beamsteering, 74, 99
- beamwidth, 47, 48
- bearing estimation, 169
- Binomial weights, 82
- bit error rate, 144
- Blackman weights, 83
- blind adaptive, 251
- blind equalization, 239
- block codes, 160
- Bohman weights, 87
- boresight, 46
- Butler matrix, 94, 95
- Capon estimate, 179–181
- Cartesian basis vectors, 173
- Central Limit Theorem, 110, 131
- channel bandwidth, 129, 152
- channel coding, 159, 160
- channel impulse response, 141
- channel matrix, 161, 162
- characteristic determinant, 175
- circular array, 89–91
- Clarke flat fading, 138, 139
- column vector, 170
- conjugate gradient, 227, 246–251
- conjugate matched arrays, 98
- constant modulus (CM), 238–246
- constant modulus algorithm (CMA), 227, 240
- convergence, 210, 229, 232
- convolutional codes, 160
- correlation matrix, 119, 120; 175–178
- cost function, 2, 219, 225
- covariance matrix (*see also* correlation matrix), 177
- autocorrelation, 117, 118
- autocorrelation-ergodic, 118
- Bartlett estimate, 178, 179
- Bartlett weights, 87
- beam solid angle, 49
- beamsteered arrays, 207
- beamsteered, 74
- beamsteering, 74, 99
- beamwidth, 47, 48
- bearing estimation, 169
- Binomial weights, 82
- bit error rate, 144
- Blackman weights, 83
- blind adaptive, 251
- blind equalization, 239
- block codes, 160
- Bohman weights, 87
- boresight, 46
- Butler matrix, 94, 95
- Capon estimate, 179–181
- Cartesian basis vectors, 173
- Central Limit Theorem, 110, 131
- channel bandwidth, 129, 152
- channel coding, 159, 160
- channel impulse response, 141
- channel matrix, 161, 162
- characteristic determinant, 175
- circular array, 89–91
- Clarke flat fading, 138, 139
- column vector, 170
- conjugate gradient, 227, 246–251
- conjugate matched arrays, 98
- constant modulus (CM), 238–246
- constant modulus algorithm (CMA), 227, 240
- convergence, 210, 229, 232
- convolutional codes, 160
- correlation matrix, 119, 120; 175–178
- cost function, 2, 219, 225
- covariance matrix (*see also* correlation matrix), 177

- D-BLAST, 163
- delay
 - excess, 142
 - maximum excess, 143
 - mean excess, 143
 - spread, 143
- delay spread, 143
- determinant, 171
- diffraction, 128
- digital beamforming, 89, 208
- dipole, 53, 55
- direction finding, 4, 5
- direction of arrival, 169
- directivity, 48
- dispersion, 151
- diversity, 158
- diversity, 158
- Dolph-Chebyshev weights, 87
- Doppler spread, 155
- Doppler, 137

- effective aperture, 49
- eigenvalues, 174
- eigenvectors, 174
- electronic beamsteering, 207–208
- element factor, 67
- equalization, 156, 157
- ergodic capacity, 163
- ergodic process, 117
- ergodicity, 115–118
 - autocorrelation-ergodic, 118
 - mean-ergodic, 116
 - variance-ergodic, 117
- Erlang distribution, 113, 132
- error correction codes, 159, 160
- error detection codes, 159, 160
- ESPRIT estimate, 197–201
- even array, 80
- excess delay, 142
- expectation, 108, 109
- Exponential distribution, 113, 132

- Faraday rotation, 123
- Faraday, M., 9
- far-field, 38
- fast fading, 129
- finite length dipole, 55
- fixed beam array, 93
- fixed beam, 1, 209, 227
- flat earth model, 124
- flat fading, 129
- flat-earth model, 28–32; 124–127

- forgetting factor, 235–237
- FPGA, 3
- Fraunhofer region, 38
- frequency-selective fading, 129, 155
- Fresnel coefficients, 24
- Fresnel region, 38
- Friis formula, 50, 51
- Friis, H.T., 50

- gain, 49
- Gauss, C. F., 9
- Gaussian density, 110
- Gaussian distribution, 110
- Gaussian weights, 84
- geolocation, 4
- gradient based algorithms, 227

- Hamming weights, 83
- Hanning weights, 87
- Helmholtz wave equation, 11, 12
- Helmholtz, H., 11
- Hermitian transpose, 170

- identity matrix, 173
- intersymbol interference, 155
- isotropic, 39–41

- Kaiser-Bessel weights, 85
- Keller's laws of diffraction, 128
- knife-edge diffraction, 31, 32

- Laplace distribution, 115, 150
- least mean square (LMS), 209, 227–231
- least squares constant modulus algorithm (LS-CMA), 240–246
- linear antennas, 53–58
- linear array, 65–79
- Linear Prediction estimate, 181, 182
- Log-normal distribution, 129, 154
- loop antenna, 58–61

- mainlobe, 44
- MANET, 3
- matrix inverse, 174
- maximum arrival angle, 145
- Maximum Entropy estimate, 182, 183
- maximum excess delay, 143
- maximum likelihood (ML), 221–223
- maximum signal to interference ratio, 211–218
- Maxwell, J. C., 9

- Maxwell's equations, 9–19
- mean arrival angle, 145
- mean excess delay, 143
- mean-square error (MSE), 207, 218–221
- MIMO, 3, 160–163
- minimum mean-square error (MMSE), 218–221
- minimum variance distortionless response (MVDR), 223–226
- Min-Norm estimate, 185–187
- moment generating function, 109, 147
- Multipath Intensity Profile (MIP), 142
- multipath intensity profile, 142
- multipath power gain, 143, 154
- multipath, 127
- Multiple Input Multiple Output (MIMO), 3, 160–163
- MUSIC estimate, 187–191

- near-field, 38
- noise subspace, 178
- Nuttall weights, 80, 87

- odd array, 81
- outage probability, 133

- Pascal's triangle, 82
- path gain factor, 29–33
- pattern multiplication, 67
- pattern, 44
- periodogram, 169, 178
- phase constant, 11
- phase locked loop (PLL), 209
- phased arrays, 207
- Pisarenko Harmonic Decomposition estimate, 183–185
- planar array, 91–93
- polarization diversity, 158
- power angular profile (PAP), 145
- power delay angular profile (PDAP), 151
- power delay profile (PDP), 142
- power density, 39–42
- power pattern, 44
- power spectral density, 117, 145
- Poynting, J.H., 40
- probability density functions, 106, 107
 - Erlang, 113
 - exponential, 113
 - Gaussian, 110
 - Laplace, 115
- Rayleigh, 111
- Rician, 114
- uniform, 111
- profile
 - power angular profile (PAP), 145
 - power delay angular profile (PDAP), 151
 - power delay profile (PDP), 142
- Propagation channel, 123–163
- propagation constant, 11
- propagation, 11–14; 123–163
- pseudospectrum, 178

- radiation intensity, 42–44
- radiation pattern, 38
- RAKE receiver, 158
- Rayleigh distribution, 111, 131
- Rayleigh fading, 129
- Rayleigh scattering, 128
- recursive least squares (RLS), 234–238
- reflection, 128
- reflection, 21–27
- refraction, 128
- regions
 - antenna, 38
 - far-field, 38
 - Fraunhofer, 38
 - Fresnel, 38
 - near-field, 38
- restrodirective arrays, 98–101
- Riccati equation, 237
- Rician distribution, 114, 135
- Rician factor, 135
- root-Min-Norm, 195
- root-MUSIC, 191–195
- row vector, 170

- sample matrix inversion (SMI), 231–234
- scanning, 209
- scattering, 128
- SDMA, 5, 93, 207, 252
- self-focusing arrays, 98
- self-phasing arrays, 98
- shading (*see also* weighting), 80
- Sherman Morrison-Woodbury (SMW), 236
- sidelobe cancellers, 95–98, 209
- sidelobes, 44
- signal subspace, 178
- signal to interference ratio (SIR), 211–218
- Single Input Single Output (SISO), 160
- slow fading, 129

270 Index

- smart antenna, 1–7, 207–266
- smart antennas, 207–260
- software defined radio, 3
- spectral estimation (*see also*
angle-of-arrival estimation),
169, 201
- spherical-earth, 14, 15
- spread
 - angular, 146
 - delay, 143
 - Doppler, 155
- spreading sequence array weights,
251–260
- stationarity, 115, 116
 - strict-sense, 116
 - wide-sense, 116
- steepest descent, 209, 228, 242, 246
- step-size parameter, 228, 229
- strict sense stationary, 116
- switched beam, 1, 251, 260

- taps, 141
- time diversity, 158
- time-reversal arrays, 98
- time-reversal mirrors, 98
- trace, 173
- transmission, 21–27
- transpose, 170
- turbo codes, 160

- uniform distribution, 111, 131
- uniform array, 68

- Vandermonde, 69, 171
- variance-ergodic, 163
- V-BLAST, 163
- vector potential, 51–53

- waveform diversity, 251
- waveform diversity, 3, 251
- weighting, 79–88
- weights, 82–87, 176
 - Bartlett, 87
 - Bohman, 87
 - Binomial, 82
 - Blackman, 83
 - Blackman-Harris, 87
 - Hamming, 83
 - Hanning, 87
 - Gaussian, 84
 - Kaiser-Bessel
 - Dolph-Chebyshev, 87
 - Nuttall, 87
- wide-sense stationary, 116, 141
- Wiener-Khinchin pair, 119
- windowing (*see also* weighting), 80
- WLAN, 3
- WLL, 3
- WMAN, 3